Natura conștiinței

Creier, minte și limbaj

MAXWELL BENNETT
DANIEL DENNETT
PETER HACKER ȘI JOHN SEARLE

Tipărit în Spania

REZUMAT

Introducere 9
DANIEL ROBINSON

CONTROVERSA

Extrase din *Fundamentele filozofice ale neuroștiinței...* 15
MAXWELL BENNETT ȘI PETER HACKER
Neuroștiințe și Filosofie 69
MAXWELL BENNETT

REFUTĂRILE

Filosofia ca antropologie naivă 93
DANIEL DENNETT
Relocarea conștiinței în creier 121
IOAN SEARLE

RĂSPUNS LA REFUTĂRI

Presupunerile conceptuale ale neuroștiinței cognitive ... 159

MAXWELL BENNETT ȘI PETER HACKER

Epilogul 203
MAXWELL BENNETT

Căutarea continuă 215
DANIEL ROBINSON

Note 245

INTRODUCERE

DANIEL ROBINSON

În 2003, Blackwell a publicat lucrarea lui Max Bennett și Peter Hacker, „ *Fundamentele filosofice ale neuroștiinței ”. Cartea a atras* imediat atenția deoarece a reprezentat prima evaluare sistematică a fundamentelor conceptuale ale neuroștiinței , așa cum au fost stabilite de oamenii de știință și filosofi. La atractivitatea lucrării au contribuit două anexe dedicate unei critici detaliate a scrierilor influente ale lui John Searle și Daniel Dennett. Max Bennett, un neurocercetător desăvârșit, i-a identificat pe bună dreptate pe Searle și Dennett drept cei mai citiți filosofi din comunitatea neuroștiințelor și a dorit să le explice cititorilor de ce el și Hacker nu erau de acord cu opiniile lor.

În toamna anului 2004, comitetul de programare al Asociației Americane de Filosofie (APA) i-a invitat pe Bennett și Hacker să participe la o sesiune „Autori și critici” la reuniunea Asociației din 2005, desfășurată la New York. Alegerea criticilor nu putea fi mai bună: Daniel Dennett și John Searle au fost de acord să scrie răspunsuri la criticile pe care Bennett și Hacker le-au adus lucrărilor lor. Conținutul acestei cărți se bazează pe acea sesiune APA de trei ore. Owen Flanagan a prezidat sesiunea, care a fost marcată de un schimb de replici neobișnuit de animat între participanți. Înainte de sesiune, Dennett și Searle au furnizat fiecare versiuni scrise ale replicilor lor, la care Bennett și Hacker au răspuns apoi.

Conștientă pe deplin de importanța subiectelor filosofice discutate, Wendy Lochner, redactorul secției de filosofie de la Columbia University Press, le-a sugerat participanților ca lucrările să fie publicate sub formă de carte. O dezbatere pasionată își pierde adesea vibranța și entuziasmul în formă scrisă. Cititorul este forțat să reconstruiască evenimentul în imaginația sa, pornind de la fragmentarea și artificiul prozei editate. Cred că este corect să spunem că această carte nu suferă de această limitare comună. În eseurile și schimburile de replici prezentate aici, cititorul va recunoaște puterea motivatoare a pasiunii intelectuale. Participanții abordează subiectele lor în profunzime. Contribuțiile lor remarcabile de-a lungul mai multor decenii le dau dreptul să fie luați în serios. Mai mult, miza este mult mai mare decât este tipic în acest tip de discuție. La urma urmei, proiectul neuroștiinței cognitive nu este altceva decât încorporarea a ceea ce numim cu plăcere natura umană în însăși structura științei. Dennett și Searle, cu o încredere care ar putea fi confundată cu entuziasmul, sunt înclinați să creadă că procesul de integrare este pe drumul cel bun . Bennett și Hacker, pe de altă parte , cu o prudență care ar putea fi confundată cu scepticism, ridică posibilitatea ca proiectul în sine să se bazeze pe o neînțelegere.

Am fost onorat să primesc cererea de a scrie un capitol final pentru cartea planificată. În acesta, îmi rezum cele mai profunde opinii pe această temă, cântărind în același timp argumentele și contraargumentele abile ale principalelor figuri implicate în dezbatere. Am încredere că cititorii își vor da seama că am puține idei definitive pe această temă. Recunosc efortul hotărât al lui Searle și Dennett de a oferi un model credibil și viabil despre cum este înțeleasă viața mentală prin intermediul evenimentelor care se petrec sub piele. Norbert Wiener, unul dintre cu adevărat înțelepții oamenilor de știință, a subliniat că cel mai bun model material al unei pisici este o pisică, de preferință pisica însăși. Cu toate acestea, fără modele - chiar și cele înfrumusețate cu trăsături antropomorfe - însăși complexitatea lumii reale frustrează în cele din urmă progresul științific în orice domeniu. Nu există calcule sau ecuații care să stabilească limitele în care trebuie să fie limitată imaginația creatorului de modele.

În cele din urmă, aceste tipuri de întrebări își găsesc răspunsul în domeniul esteticii. Prin aceasta nu vreau să spun că există mai puțin loc pentru rigoarea analitică : în cel mai bun caz, analiza filosofică este o strădanie estetică. Cu siguranță, acesta este ceea ce îi atrage pe fizicieni și matematicieni spre „eleganță”. Nu este estetica cea care folosește briciul lui Occam pentru a

determina rafinamentul, măsura , proporționalitatea și coerența? Sunt convins că, tocmai conform acestor criterii, cititorii vor găsi în critica lui Bennett-Hacker - în special în critica bogată din punct de vedere filosofic și bine informată a lui Peter Hacker - nu o înclinație spre scepticism , ci o aplicare atentă și într-adevăr *elegantă* a celor mai fine instrumente concepute de filosofi.

Acestea fiind spuse, este important să mergem mai departe și să recunoaștem că viața noastră *trăită în realitate este puțin probabil* să-și dezvăluie realitatea interioară completă , schimbătoare, adesea capricioasă și minunat de complexă tabelelor de adevăr, unei mașini Turing sau unui laparoscop. Nu este surprinzător faptul că filosoful care ne-a oferit atât de des prima idee despre un subiect important este și cel care are ultimul cuvânt în această privință. Mă refer, desigur , la Aristotel. Trebuie să căutăm precizia în acele lucruri care o permit și să alegem instrumentele potrivite pentru sarcina pe care o avem. În cele din urmă, explicațiile noastre trebuie să aibă o relație inteligibilă cu ceea ce intenționăm să explicăm. Demograful care afirmă cu o acuratețe lăudabilă că familia medie este formată din 2,53 membri nu simte nevoia să ne reamintească că nu există așa ceva ca 0,53 persoane. Astfel de date nu sunt destinate să descrie natura elementelor numărate; acea cantitate este pur și simplu rezultatul lor. Ideea este, desigur, că precizia științifică sau, de altfel, precizia aritmetică, nu ne poate spune aproape nimic despre ceea ce a fost analizat cu o asemenea precizie. Aici, ca și în alte părți , maxima predominantă este *caveat emptor (ia comoara).* [atenție, cumpărător].

Cititorii vor aborda acest discurs cu un interes autentic, chiar și cu o urmă de vanitate, deoarece *îi privește.* Propriile lor criterii estetice vor prevala asupra materialelor de acest fel și, în cele din urmă, ei vor decide dacă explicațiile oferite au vreo legătură inteligibilă cu ceea ce contează cu adevărat. Însă un juriu bun trebuie să respecte dovezile disponibile, ghidat în deliberările sale de judecata solidă a probelor. Cititor răbdător! Jurat demn! Iată câteva dovezi (din neuroștiința cognitivă) și o expunere excepțional de clară a regulilor aplicabile luării lor în considerare. Nu este nevoie să ne grăbim să emitem un verdict...

DANIEL N. ROBINSON

U\ POLÉMICA

FUNDAMENTE FILOZOFICE DE NEUROȘTIINȚE

Introducere

MAXWELL BENNETT ȘI PETER HACKER

Cartea *Fundamentele filozofice ale neuroștiinței* [Fundamentele filozofice ale neuroștiinței] prezintă roadele unui efort de colaborare între un neurocercetător și un filosof. Abordează fundamentele conceptuale ale *neuroștiinței cognitive,* fundamente constituite de relațiile structurale dintre conceptele psihologice implicate în cercetarea bazelor neuronale ale capacităților cognitive, afective și volitive umane. Investigarea relațiilor logice dintre concepte este o sarcină filosofică. Ghidarea acestei cercetări pe căi specifice, astfel încât să lumineze investigarea creierului , este o sarcină neuroștiințifică. De aici și demersul nostru comun.

Dacă vrem să înțelegem structurile și dinamica neuronală care fac posibile percepția, gândirea, memoria, sentimentele și comportamentul intențional, este esențială clarificarea acestor concepte și categorii. Ambii autori, care au abordat această cercetare din perspective foarte diferite, s-au simțit nedumeriți și uneori inconfortabili în legătură cu utilizarea conceptelor psihologice în neuroștiința actuală. Această nedumerire a provenit din implicațiile a ceea ce un anumit neurocercetător a afirmat despre creier și minte sau din motivele pentru care un neurocercetător credea...

Experimentele lor au clarificat capacitatea psihologică studiată sau presupunerile conceptuale ale întrebărilor puse. Neliniștea provenea din suspiciunea că, în unele cazuri, conceptele nu erau înțelese, erau aplicate incorect sau erau utilizate dincolo de domeniul lor specific de aplicare. Și cu cât investigam mai mult, cu atât ne convingeam mai mult că, în ciuda progreselor impresionante din neuroștiința cognitivă, ceva nu era în regulă cu teoretizarea generală.

Întrebările empirice despre sistemul nervos sunt domeniul neuroștiinței. Sarcina sa este de a stabili faptele referitoare la structurile și operațiunile neuronale. Sarcina neuroștiinței *cognitive este* de a explica condițiile neuronale care fac posibile funcțiile perceptive, cognitive, cogitative, afective și volitive . Cercetarea experimentală confirmă sau contestă astfel de teorii explicative. În schimb, întrebările conceptuale (cele care, de exemplu, privesc conceptele de minte sau memorie, gândire sau imaginație), descrierea relațiilor logice dintre concepte (cum ar fi cele dintre conceptele de percepție și senzație sau conștiință și conștiință de sine) și examinarea relațiilor structurale dintre diferite câmpuri conceptuale (de exemplu, între psihologic și neuronal sau între mental și comportamental) sunt domeniul propriu al filosofiei.

Întrebările conceptuale preced întrebările despre adevăr și falsitate. Sunt întrebări care privesc *formele noastre de reprezentare,* nu adevărul sau falsitatea afirmațiilor empirice. Aceste forme sunt presupuse în afirmațiile științifice adevărate (și false) și în teoriile științifice corecte *(și* incorecte). Ele nu determină ce este empiric adevărat sau fals, ci mai degrabă ce are sens și ce nu. Prin urmare, întrebările conceptuale nu sunt relevante pentru cercetarea și experimentarea științifică, nici pentru teoretizarea științifică. Acest lucru se datorează faptului că niciuna dintre aceste investigații și teoretizări *presupune* conceptele și relațiile conceptuale în cauză. Interesul nostru aici nu este în liniile de demarcație ale breslelor, ci în distincțiile dintre tipuri logic diferite de cercetare intelectuală.

Este de o importanță capitală să se facă distincția între întrebările conceptuale și cele empirice. Atunci când o întrebare conceptuală este confundată cu una științifică, aceasta apare inevitabil extrem de abstrusă. În astfel de cazuri, se pare că știința ar trebui să poată descoperi adevărul a ceea ce este investigat prin teorie și experiment, și totuși eșuează în mod constant în această încercare. Acest lucru nu este surprinzător, deoarece întrebările conceptuale nu pot fi abordate prin metode empirice de investigare, la fel cum problemele din matematica pură nu pot fi rezolvate prin metodele fizicii. Mai mult, atunci când problemele empirice sunt abordate fără o claritate conceptuală adecvată, apar inevitabil întrebări ambigue, conducând investigația pe căi fără rezultat. Orice obscuritate cu privire la conceptele relevante se va reflecta într-o lipsă corespunzătoare de claritate a întrebărilor și, în consecință, în proiectarea experimentelor menite să le răspundă. Și este

de așteptat ca orice inconsecvență în înțelegerea structurii conceptuale relevante să se manifeste în inconsecvențe în interpretarea rezultatelor experimentelor.

Neuroștiința cognitivă operează la granița dintre două domenii, neurofiziologia și psihologia, ale căror concepte sunt categoric diferite. Relațiile logice sau conceptuale dintre fiziologic și psihologic sunt problematice. Numeroase concepte și categorii psihologice sunt dificil de definit clar. Relațiile dintre minte și creier și dintre psihologic și comportamental sunt derutante. Stranietatea în fața acestor concepte și a expresiilor lor, precum și a acestor aparente „domenii" și a relațiilor dintre ele, a fost o caracteristică [a] neurofiziologiei încă de la originile sale.2 În ciuda marilor progrese în neuroștiință de la începutul secolului al XX-lea, conduse de Charles Sherrington, bateria de întrebări conceptuale cunoscută popular sub numele de problema minte-corp sau minte-creier a rămas la fel de dificil de rezolvat ca întotdeauna, după cum o demonstrează ideile carteziene eronate susținute de Sherrington și unii dintre colegii și protejații săi, precum Edgar Adrian, John Eccles și Wilder Penfield. În ciuda strălucirii incontestabile a muncii lor, au persistat confuzii conceptuale serioase. [3] Scopul cercetării noastre din această carte este de a determina dacă generația actuală de neurocercetători a depășit cu succes confuziile conceptuale ale generațiilor anterioare sau doar a înlocuit un amestec conceptual cu altul.

O astfel de confuzie este evidentă în atribuirea persistentă de atribute psihologice creierului. În timp ce Sherrington și protejații săi atribuiau atribute psihologice minții (concepută ca o substanță particulară, poate imaterială, distinctă de creier), neurologii actuali tind să atribuie același set de atribute psihologice creierului (care este considerat în mod obișnuit, deși nu uniform, a fi identic cu mintea). Dar mintea, susținem noi, nu [este nici o substanță distinctă de creier , nici] o substanță identică cu creierul. Și demonstrăm că atribuirea de atribute psihologice creierului este incoerentă. Ființele umane posedă o mare varietate de capacități psihologice, care sunt exercitate în circumstanțele vieții atunci când percepem, gândim și raționăm, simțim emoții, dorim lucruri, facem planuri și luăm decizii. Posesia și exercitarea acestor capacități ne definesc ca genul de animal care suntem. Putem investiga condițiile neuronale și corelațiile unei astfel de posesii și exerciții. Aceasta este sarcina neuroștiinței, care descoperă din ce în ce mai multe despre toate acestea. Însă descoperirile sale nu afectează în niciun fel adevărul conceptual conform căruia aceste capacități și exercitarea lor în percepție, gândire și sentiment *sunt atribute ale ființelor umane,* nu ale părților lor, în mod specific, *nu ale creierului lor.* Ființa umană este o unitate psihofizică, un animal care poate percepe, acționa intenționat, raționa și simți emoții, un animal care folosește limbajul, care nu este doar - conștient, ci și conștient de sine; nu un creier încorporat în craniul unui corp. Sherrington,

Eccles și Penfield au conceput ființele umane ca animale a căror minte, pe care au imaginat-o ca purtătoare de atribute psihologice, este legată de creier. A presupune că creierul este purtătorul de atribute psihologice nu reprezintă niciun progres față de această concepție falsă.

Printre neurologii de astăzi, este foarte comun să se vorbească despre modul în care creierul percepe, gândește, ghicește sau crede sau despre cum una dintre emisferele sale știe lucruri pe care cealaltă nu le știe. Justificarea este de obicei că acest tip de discurs este doar o figură de stil. Dar aceasta este o greșeală gravă, deoarece forma caracteristică de explicație în neuroștiința cognitivă contemporană constă în atribuirea de atribute psihologice creierului și părților sale *pentru a explica* deținerea atributelor psihologice și exercitarea (și deficiențele în exercitarea) abilităților cognitive de către ființele umane.

Atribuirea de atribute psihologice, în special cele cognitive și cogitative, creierului este, de asemenea, o sursă de confuzie suplimentară serioasă, așa cum am demonstrat. Neuroștiința poate investiga condițiile și corelațiile neuronale ale dobândirii, deținerii și exercitării capacităților senzoriale de către oameni.

Yo

Animale. Puteți descoperi cerințele neuronale pentru exercitarea capacităților distinct umane de gândire și raționament, memorie declarativă și imaginație, emoție și voință. Puteți face toate acestea prin corelație inductivă răbdătoare între fenomenele neuronale și posedarea și exercitarea capacităților psihologice și între leziunile neuronale și -deficiențele funcțiilor mentale normale. Ceea ce *nu puteți* face este *să înlocuiți* explicațiile neurologice cu marea varietate de explicații psihologice comune ale activităților umane în termeni de motive, intenții, scopuri, obiective, valori, norme și convenții . Și [nu puteți] explica *modul* în care un animal percepe sau gândește referindu-vă la percepția sau gândirea creierului sau a oricăreia dintre părțile sale. Nu are sens să atribuim astfel de atribute psihologice altceva decât animalului ca întreg. Animalul este cel care percepe, nu părți ale creierului său, iar ființele umane sunt cele care gândesc și raționează, nu creierul lor. Creierul și activitățile sale ne *permit nouă ,* nu lui, să percepem și să gândim, să simțim emoții și să dezvoltăm proiecte pe care încercăm să le transformăm în realitate.

Dacă răspunsul imediat al multor neurocercetători la acuzația de confuzie conceptuală este de a afirma că atribuirea predicatelor psihologice creierului este doar

> o figură de stil, reacția lor la faptul demonstrabil că teoriile lor explicative atribuie capacități psihologice creierului *într-un mod departe de a fi trivial* este de a sublinia că o astfel de eroare este inevitabilă din cauza deficiențelor limbajului. Discutăm această idee falsă7 [și] demonstrăm că marile descoperiri ale neuroștiinței *nu necesită* această formă falsă de explicație, că ceea ce a fost descoperit poate fi perfect descris și explicat cu limbajul pe care îl avem la dispoziție. Și demonstrăm acest lucru citând fenomenele mult dezbătute rezultate dintr-o comisurotomie, descrise (în opinia noastră, descrise greșit) de Sperry, Gazzaniga și alții.8

În partea a doua a lucrării *Fundamentele filozofice ale neuroștiinței* (denumit în continuare PFN) Investigăm utilizarea conceptelor de percepție , memorie, imagerie mentală, emoție și voință în teoretizarea actuală a neuroștiinței. De la caz la caz, demonstrăm că lipsa de claritate conceptuală - incapacitatea de a acorda o atenție adecvată structurii conceptuale relevante - a fost adesea o sursă de erori teoretice și baza unor inferențe eronate. Este o greșeală, o eroare conceptuală, să presupunem că percepția este o chestiune de a înțelege o *imagine* în minte (Crick, Damasio, Edelman), de a formula o *ipoteză* (Helmholtz, Gregory) sau de a genera *un model descriptiv în trei dimensiuni* (Marr). Este confuz, o confuzie conceptuală, să încadrăm problema legării *ca* o combinare a datelor de formă, culoare și mișcare pentru a forma *imaginea* obiectului perceput (Crick, Kandel, Wurtz). Este inexact, inexact din punct de vedere conceptual, să presupunem că memoria este întotdeauna o memorie a trecutului sau să credem că amintirile pot *fi stocate* în creier sub forma puterii conexiunilor sinaptice (Kandel, Squire, Bennett). Și este o greșeală, o greșeală conceptuală, să presupunem că investigarea setei, foamei și dorinței sexuale este o investigație a emoțiilor (Roluri) sau să credem că funcția emoțiilor este de a ne informa despre starea noastră viscerală și musculo-scheletică (Damasio).

Reacția imediată la astfel de observații critice ar putea fi una de indignare și neîncredere. Cum poate o știință înfloritoare să fie fundamental greșită? Cum ar putea exista o confuzie conceptuală inevitabilă într-o știință bine stabilită? Dacă există concepte problematice, acestea pot fi cu siguranță ușor înlocuite cu altele care nu sunt problematice și care servesc acelorași scopuri explicative . Astfel de răspunsuri dezvăluie o înțelegere deficitară a relației dintre forma de reprezentare și faptele reprezentate și o neînțelegere a naturii erorii conceptuale. De asemenea, ele dezvăluie o ignoranță a istoriei științei în general și a neuroștiinței în special.

Știința nu este mai imună la erori și confuzii conceptuale decât orice altă activitate

intelectuală. Istoria științei este plină de rămășițe ale unor teorii care nu erau doar incorecte din punct de vedere factual , ci și eronate din punct de vedere conceptual. Teoria combustiei a lui Stahl, de exemplu, conținea eroarea *conceptuală de a atribui, în anumite circumstanțe, o pondere negativă flogistonului - o idee care nu avea sens în cadrul său conceptual, cel al* fizicii newtoniene . Celebrele critici ale lui Einstein la adresa teoriei eterului (presupusul mediu prin care se credea că se propagă lumina) au fost îndreptate nu numai către rezultatele experimentelor Michelson-Morley, care nu reușiseră să detecteze niciun efect [de] mișcare absolută, ci și către confuzia conceptuală privind mișcarea relativă inerentă rolului atribuit eterului în explicarea inducției electromagnetice. Neuroștiința nu a făcut excepție, așa cum arătăm în studiul nostru istoric.9 Nu există nicio îndoială că este o *știință înfloritoare astăzi,* dar acest lucru nu o face imună la confuzii și încurcături conceptuale . Cinematica newtoniană a fost o știință înfloritoare, dar acest lucru nu a împiedicat confuziile conceptuale ale lui Newton cu privire la inteligibilitatea acțiunii la distanță și nici nedumerirea sa (nerezolvată până la Hertz) cu privire la natura forței. În mod similar, așadar, cea mai mare realizare a lui Sherrington - explicarea acțiunii integrative a sinapselor din măduva spinării și, în consecință, eliminarea odată pentru totdeauna a ideii confuze de „suflet spinal" - a fost perfect compatibilă cu confuziile conceptuale privind „sufletul cerebral" sau mintea și relația sa cu creierul. În mod similar, realizările extraordinare ale lui Penfield în identificarea locațiilor funcționale din cortex, precum și în dezvoltarea unor tehnici neurochirurgicale strălucite, au fost perfect compatibile cu confuziile importante despre relația dintre minte și creier și despre „funcția superioară a creierului" (o idee pe care a împrumutat-o de la Hughlings Jackson).

Pe scurt, confuzia conceptuală *poate* coexista cu știința înfloritoare. Acest lucru poate părea ciudat. Dacă știința poate înflori în ciuda unei astfel de confuzii conceptuale, de ce și-ar face oamenii de știință griji în privința asta? Recifele ascunse nu înseamnă că mările sunt nenavigabile; ele înseamnă doar că sunt periculoase. Întrebarea discutabilă este cum se manifestă progresul peste aceste recife. Confuziile conceptuale pot apărea în moduri diferite și în diferite etape ale cercetării. În unele cazuri, lipsa clarității conceptuale poate să nu afecteze plauzibilitatea întrebărilor sau valoarea experimentelor, ci doar interpretarea rezultatelor acestora și implicațiile lor teoretice. De exemplu, Newton a început *Optica* în căutarea unor perspective asupra naturii culorii. Cercetările sale au reprezentat o - contribuție durabilă la știință. Dar concluzia sa că „culorile sunt senzații în senzoriu"

demonstrează că nu a atins genul de înțelegere pe care îl căuta. Deoarece culorile, oricare ar fi ele, nu sunt „senzații în senzoriu". Astfel, Newton, în măsura în care era preocupat de *interpretarea* rezultatelor cercetărilor sale, avea motive întemeiate să fie preocupat de confuziile conceptuale cu care lucra, întrucât acestea îi împiedicau calea către o interpretare corectă.

Totuși, în alte cazuri, confuzia conceptuală afectează și cercetarea empirică. Este foarte posibil ca adresarea unor întrebări greșite să trivializeze cercetarea.10 Iar la cealaltă extremă, construcția defectuoasă a conceptelor și a structurilor conceptuale va duce, în unele cazuri, la cercetări care, deși nu sunt triviale, vor fi incapabile să demonstreze ceea ce au fost concepute să demonstreze.11 În aceste cazuri, știința poate să nu fie atât de înfloritoare pe cât pare, iar cercetarea conceptuală poate fi necesară pentru a identifica și elimina problemele.

Sunt aceste confuzii conceptuale *inevitabile* ? Deloc. Singurul scop al scrierii acestei cărți a fost acela de a arăta cum să le evităm. Dar, desigur, nu pot fi evitate dacă totul rămâne neschimbat. *Pot fi* evitate, dar dacă sunt evitate, anumite tipuri de întrebări vor înceta să fie puse , deoarece va deveni clar că se bazează pe o neînțelegere. După cum spune Hertz în magnifica sa introducere la *Principiile mecanicii,* „Când aceste contradicții nefericite vor fi eliminate, [...] mintea noastră, acum liberă de discordie , va înceta să formuleze întrebări false." În mod similar, anumite tipuri de inferențe vor înceta să fie trase dintr-un anumit corp de studii empirice, deoarece se va dezvălui că au puțin sau nimic de-a face cu problema pe care au fost menite să o clarifice, chiar dacă pot avea legătură cu *altceva.*

Dacă există concepte problematice, pot fi acestea înlocuite cu altele care îndeplinesc aceeași funcție explicativă? Oamenii de știință sunt întotdeauna liberi să introducă concepte noi dacă le consideră pe cele existente inadecvate sau insuficient dezvoltate. Însă preocuparea noastră în această carte nu constă în utilizarea de noi concepte tehnice, ci în utilizarea greșită a conceptelor vechi, non-tehnice: minte și corp, gândire și imaginație, senzație și percepție, cunoaștere și memorie, mișcare voluntară și conștiință și conștiință de sine. Nu este nimic inadecvat în aceste concepte în ceea ce privește scopurile pe care le servesc. Nu există niciun motiv să credem că ar trebui înlocuite în contextele pe care le luăm în considerare. Problema constă în interpretarea eronată a acestor concepte de către neurocercetători și în neînțelegerile care rezultă din aceasta. Astfel de neînțelegeri pot fi remediate printr-o explicație corectă a categoriei

logico-gramaticale a conceptelor în cauză. Și exact asta am încercat să oferim.

Dacă admitem că neurologii pot să nu folosească aceste concepte comune în același mod în care o face o persoană obișnuită, cu ce drept poate filosofia să pretindă să le corecteze? Cum poate filosofia să pună la îndoială cu atâta încredere claritatea și coerența conceptelor folosite de oamenii de știință competenți? Cum poate fi filosofia în măsură să proclame că anumite afirmații ale unor neurologi sofisticați *sunt lipsite de sens?* În paginile următoare, vom rezolva aceste îndoieli metodologice. Dar poate că o clarificare inițială asupra acestui punct va risipi unele dintre ele. Adevărul și falsitatea sunt pentru știință ceea ce sensul și nonsensul sunt pentru filosofie. Erorile observaționale și teoretice duc la falsitate; erorile conceptuale, la lipsă de sens. Cum pot fi investigate limitele sensului? Numai examinând utilizarea cuvintelor. Nonsensul apare adesea atunci când o expresie este folosită contrar regulilor care guvernează utilizarea sa. Expresia în cauză poate fi o expresie comună, netehnică, caz în care regulile de utilizare a acesteia pot fi deduse din utilizarea sa obișnuită și din explicațiile standard ale sensului său. Sau poate fi un termen tehnic dintr-o disciplină, caz în care regulile de utilizare a acestuia trebuie deduse din introducerea termenului de către teoretician și din explicațiile pe care le oferă cu privire la utilizarea sa stipulată. Ambele tipuri de termeni pot fi utilizate greșit, iar când se întâmplă acest lucru, apare un nonsens - o combinație de cuvinte care este exclusă din limbaj. Acest lucru se datorează faptului că fie nu s-a stipulat nimic despre ce înseamnă termenul în contextul aberant în cauză, fie această combinație de cuvinte este de fapt exclusă de o regulă care specifică că nu există așa ceva (de exemplu, că nu există așa ceva ca „la est de Polul Nord"), că aceasta este o combinație de cuvinte care nu are nicio utilitate. Nonsensul apare frecvent și atunci când unei expresii i se dă o utilizare nouă, poate tehnică sau cvasi-tehnică, iar noua utilizare se suprapune în mod accidental cu cea veche - de exemplu, atunci când se trag inferențe din propoziții care conțin noul termen și care sunt legitime doar în raport cu cel vechi. Identificarea unor astfel de încălcări ale limitelor sensului este sarcina criticului conceptual. Desigur, nu este suficient să se demonstreze că un anumit om de știință a folosit un termen contrar utilizării sale obișnuite, deoarece este foarte posibil ca acesta să îl folosească într-un sens nou . Criticul trebuie să demonstreze că omul de știință intenționează să folosească termenul în sensul său obișnuit, dar nu o face; sau că intenționează să îl folosească într-un sens nou, dar fără să realizeze că a suprapus noul cu vechiul sens. Ori de câte ori este posibil, omul de știință greșit îndrumat va fi judecat după propriile cuvinte. [12]

Ultima greșeală împotriva căreia dorim să avertizăm este ideea că reflecțiile

noastre sunt în mod iremediabil negative. S-ar putea presupune că singurul nostru scop este să dezaprobăm. Munca noastră poate părea, dintr-o perspectivă superficială, o strădanie pur distructivă , care nu oferă nici ajutor, nici noi căi de urmat. Mai rău, ar putea părea că încercăm să confruntăm filosofia și neuroștiința cognitivă. *Nimic nu ar putea fi mai departe de adevăr.*

Am scris această carte din admirație pentru realizările neuroștiinței secolului XX și din dorința de a contribui la această disciplină . Însă *singurele* modalități prin care cercetarea conceptuală poate ajuta o disciplină empirică sunt prin identificarea erorii conceptuale (dacă există) și prin furnizarea unei hărți care să ajute cercetătorii empirici să nu se abate de la calea cea dreaptă a sensului . Astfel, fiecare dintre investigațiile noastre are două aspecte . Pe de o parte, am încercat să identificăm problemele și încurcăturile conceptuale prezente în teoriile actuale importante ale percepției, memoriei, imaginației, emoției și voinței. Mai mult, susținem că o mare parte din scrierile despre natura conștiinței și a conștiinței de sine sunt pline de dificultăți conceptuale. Acest aspect al investigațiilor noastre este, trebuie să recunoaștem, negativ și critic. Pe de altă parte, ne-am străduit, de la caz la caz, să oferim o reprezentare clară a câmpului conceptual al fiecăruia dintre conceptele problematice. Acesta este un demers constructiv . Avem încredere că aceste prezentări generale conceptuale îi vor ajuta pe neurocercetători în reflecțiile lor pre-experimentale privind designul. Totuși, sarcina cercetării conceptuale nu poate fi de a formula ipoteze empirice care să poată rezolva problemele empirice cu care se confruntă oamenii de știință. A deplânge faptul că cercetarea filosofică asupra neuroștiinței cognitive nu a produs o nouă teorie neuroștiințifică este ca și cum te-ai plânge unui matematician care a demonstrat o nouă teoremă că nu este o nouă teorie fizică.

FUNDAMENTE FILOZOFICE DE NEUROȘTIINȚE

Extras din capitolul 3

MAXWELL BENNETT și PETER HACKER

3.1 CONFUZII MEROLOGICE ÎN NEUROȘTIINȚA COGNITIVĂ

Atribuirea atributelor psihologice creierului

Figurile de frunte ale primelor două generații de neurocercetători moderni au fost fundamental carteziene. La fel ca Descartes, ei au distins mintea de creier și i-au atribuit atribute psihologice. Prin urmare, atribuirea unor astfel de predicate ființelor umane a fost derivată, ca în metafizica cartesiană. Cu toate acestea, a treia generație de neurocercetători a respins dualismul mentorilor lor. În procesul de explicare a posesiunii atributelor psihologice de către ființele umane, aceștia au atribuit astfel de atribute nu minții, ci creierului sau părților sale.

Neurologii presupun că creierul are o gamă largă de capacități cognitive, perceptive și volitive . Francis Crick afirmă:

> Ceea ce vezi nu este *de fapt* ceea ce este acolo; este ceea ce creierul tău *crede* că este acolo [...] Creierul face cea mai bună interpretare pe care o poate în funcție de

Cu experiența sa anterioară și informațiile limitate și ambigue furnizate de ochi [...] Creierul combină informațiile oferite de caracteristicile foarte diferite ale scenei vizuale (aspecte de formă, culoare, mișcare etc.) și se stabilește asupra celei mai plauzibile interpretări a tuturor acestor indicii luate împreună. [...] Ceea ce trebuie să construiască creierul este o interpretare pe mai multe niveluri a scenei vizuale. [...] [Acest proces de reconstrucție] permite creierului să conjectureze o imagine completă doar din informații parțiale - o abilitate foarte utilă. [1]

Astfel, creierul *are experiențe, crede* lucruri, *interpretează* indicii *din informațiile* furnizate și face *conjecturi.* Gerald Edelman susține că structurile interne ale creierului „categorizează, discern și recombină diversele activități cerebrale care au loc în diferite tipuri de mapare globală" și că creierul „relaționează recursiv secvențele semantice cu cele fonologice și apoi generează corespondențe sintactice, nu din reguli preexistente, ci tratând regulile generate în memorie ca obiecte ale manipulării conceptuale".[2] Prin urmare, creierul *categorizează, în esență* , „își categorizează propriile activități (în special categorizările perceptive)" și *manipulează conceptual regulile.* Colin Blakemore susține:

> Se pare că suntem conduși să spunem că astfel de neuroni [deoarece reacționează într-un mod foarte specific la, de exemplu, orientarea liniară] posedă cunoștințe. Au inteligență, deoarece sunt capabili să calculeze probabilitatea evenimentelor externe, evenimente care sunt importante pentru animalul în cauză. Iar creierul își dobândește cunoștințele printr-un proces analog raționamentului inductiv al metodei științifice clasice. Neuronii prezintă creierului argumente bazate pe caracteristicile concrete pe care le detectează, argumente pe [baza] cărora creierul își construiește ipotezele de percepție.3

Astfel, creierul *cunoaște* lucruri, *raționează* inductiv, *construiește ipoteze* bazate pe argumente, iar neuronii care îl constituie sunt *inteligenți,* capabili să *calculeze probabilități* și *să ofere argumente.* JZ Young împărtășește în mare măsură această opinie. El afirmă că „putem considera fiecare act de a vedea ca o căutare continuă de răspunsuri la întrebările

formulate de creier. Semnalele care vin de la retină constituie «mesaje» care transmit aceste răspunsuri. Creierul folosește apoi aceste informații pentru a construi o ipoteză adecvată despre ceea ce există acolo.” ^[4] În consecință , creierul *formulează întrebări, caută răspunsuri* și *construiește ipoteze.* Antonio Damasio afirmă că „creierul nostru ia de obicei decizii bune, în câteva secunde sau câteva minute, în funcție de intervalul de timp pe care îl stabilim ca fiind potrivit pentru obiectivul pe care dorim să îl atingem și, dacă poate face așa ceva, trebuie să îndeplinească o astfel de sarcină minunată cu ceva mai mult decât rațiune pură”, [5] iar Benjamin Libet subliniază că „creierul «decide» să inițieze sau cel puțin să pregătească începutul actului, înainte de a exista o conștientizare subiectivă notabilă că o astfel de decizie a avut loc.” [6] Cu alte cuvinte, creierul *decide,* sau cel puțin „decide” și *inițiază acțiunea voluntară.*

Psihologii sunt de acord. J.P. Frisby susține că „în creier trebuie să existe o descriere simbolică a lumii exterioare, o descriere proiectată pe simboluri care reprezintă diversele aspecte ale lumii de care văzul ne face conștienți”.7 Așadar , există *simboluri în creier,* iar acesta *folosește simboluri* și, probabil , *le înțelege.* Richard Gregory înțelege vederea ca fiind „probabil cea mai complexă dintre toate activitățile creierului : se bazează pe depozitele sale de date din memorie; necesită clasificări subtile , comparații și decizii logice pentru ca datele senzoriale să devină percepție” .[8] Prin urmare, creierul *vede, face clasificări, comparații* și *decizii.* Iar specialiștii în științe cognitive sunt de acord. David Marr susține că „ creierul nostru trebuie să fie cumva capabil să *reprezinte [...] informații* [...] Prin urmare, studiul vederii trebuie să includă și [...] o investigație a naturii *reprezentărilor interne prin* care *captăm aceste informații* și *le facem accesibile ca bază pentru deciziile* despre gândurile și acțiunile noastre”.^[9] Iar Philip Johnson-Laird sugerează că creierul „are acces la un model parțial al propriilor capacități” și posedă „mașinăria recursivă pentru inserarea de modele în cadrul modelelor”; conștiința, susține el, „este proprietatea unei clase de algoritmi paraleli ”. ¹⁰

Punerea la îndoială a inteligibilității atribuirii atributelor psihologice creierului

Având în vedere acest consens larg asupra modului corect de a gândi despre funcțiile creierului și despre explicarea cerințelor cauzale pentru ca oamenii să posede și să își exercite capacitățile naturale de gândire și percepție, suntem înclinați să fim copleșiți de proclamații entuziaste - despre noi domenii de cunoaștere cucerite, noi mistere dezvăluite. [Dar] trebuie să luăm lucrurile încet și să ne oprim pentru a gândi. Știm cum este pentru ființele umane să experimenteze lucruri, să vadă lucruri, să știe sau să creadă lucruri, să ia decizii, să interpreteze date echivoce, să facă conjecturi și să formuleze ipoteze. Înțelegem ce înseamnă pentru ființele umane să raționeze inductiv, să calculeze probabilități , să prezinte argumente și să clasifice și să categorizeze lucrurile pe care le întâlnesc în experiența lor. Punem întrebări și căutăm răspunsuri, folosind simbolismul - limbajul nostru - în ai cărui termeni ne reprezentăm lucrurile. Dar știm cum este pentru *creier* să vadă sau să audă, să aibă experiențe, să știe sau să creadă ceva? Avem vreo idee cum ar fi pentru *creier* să ia o decizie? Avem vreo intuiție despre cum este pentru creier (ca să nu mai vorbim de un neuron) *să raționeze* (inductiv sau deductiv), *să calculeze probabilități, să prezinte argumente, să interpreteze date* și *să formuleze ipoteze* bazate pe interpretările lor? Putem observa dacă o persoană vede un lucru sau altul; îi observăm comportamentul și îi punem întrebări. Dar cum ar fi să observăm dacă un creier vede ceva, spre deosebire de observarea creierului *unei persoane* care vede ceva? Recunoaștem când o persoană pune o întrebare și când o alta răspunde. Dar avem vreo idee cum ar fi pentru un creier să pună o întrebare sau să răspundă la ea? Acestea sunt toate atribute ale ființelor umane. Este o *descoperire nouă* faptul că și creierul desfășoară aceste tipuri de activități umane ? Sau este o inovație lingvistică, introdusă de neurocercetători, psihologi și specialiști în științe cognitive, care extinde utilizarea comună a acestor expresii psihologice din motive teoretice solide? Sau, mai tulburător, este o confuzie conceptuală? Se poate oare ca pur și simplu creierul să nu gândească sau să nu știe, să nu vadă sau să audă, să nu creadă sau să nu ghicească, să nu posede și să nu folosească informații , să nu formuleze ipoteze - cu alte cuvinte, ca aceste moduri de combinare a cuvintelor să fie lipsite de sens? Dar dacă nu este așa, de ce atâția oameni de știință distinși au crezut că aceste expresii, folosite în acest

fel, au sens?

Întrebarea dacă atributele psihologice pot fi atribuite în mod inteligibil creierului este o întrebare filosofică și, prin urmare, conceptuală, nu științifică.

Întrebarea cu care ne confruntăm este una filosofică, nu una științifică. Necesită o clarificare conceptuală, nu o investigație experimentală. Nu putem investiga experimental dacă creierul gândește sau nu, crede sau nu, ghicește sau nu, raționează sau nu, formulează ipoteze sau nu etc., până când nu știm ce ar însemna pentru creier să facă astfel de lucruri - adică, până când nu ne este clar sensul acestor expresii și nu știm când (dacă există ceva) se *consideră că* un creier face toate acestea și ce fel de dovezi susțin atribuirea unor astfel de atribute creierului. (Nu putem căuta polii Pământului până când nu știm ce este un pol - adică ce înseamnă termenul „pol" - și, de asemenea, când se consideră că un pol al Pământului a fost descoperit. Altfel, cineva s-ar putea îmbarca, precum Winnie-the-Pooh, într-o expediție la Polul Est.) Întrebarea în dezbatere este: Are sens să atribuim astfel de atribute creierului? Există cu adevărat gândirea, credința etc. a creierului ? (Există Polul Est?)

În *Investigații filozofice,* Wittgenstein face o observație profundă, direct legată de subiectul nostru: *„Numai despre ființele umane și despre lucrurile care seamănă cu ființele umane (se comportă ca ființele umane) putem spune: au senzații; văd, sunt orbi; aud, sunt surzi ; sunt conştienți sau inconștienți."* [2] Aici se află chintesența concluziilor la care vom ajunge în investigația noastră. Având în vedere laconicismul obișnuit al autorului, este necesar să detaliem această idee și să-i elucidăm ramificațiile.

Nu este un fapt evident. Nu este o certitudine că doar ființele umane și lucrurile care se comportă ca oamenii sunt subiectele acestor predicate psihologice. Dacă ar fi așa, ar fi cu siguranță o descoperire recentă a neurocercetărilor că și creierele văd și aud, gândesc și cred, formulează și răspund la întrebări. O astfel de descoperire ar demonstra, fără îndoială, că aceste lucruri pot fi spuse nu numai despre ființele umane și lucrurile care se comportă ca oamenii. Ar fi uimitor și am dori să aflăm mai multe

despre asta. Am dori să știm care ar fi dovezile pentru o astfel de descoperire remarcabilă . Dar, evident, nu este cazul. Atribuirea de atribute psihologice creierului nu este susținută de nicio descoperire neuroștiințifică care să demonstreze că, contrar convingerilor noastre anterioare, creierele gândesc și raționează de fapt, la fel cum facem și noi înșine. Neurocercetătorii , psihologii și specialiștii în științe cognitive care adoptă aceste forme de descriere nu o fac ca urmare a *observațiilor* care demonstrează că creierul gândește și raționează. Susan Savage-Rambaugh a adunat dovezi concludente care arată că cimpanzeii bonobo, atunci când sunt antrenați și învățați corespunzător , pot pune și răspunde la întrebări, pot raționa într-un mod rudimentar, pot da și respecta comenzi și așa mai departe. Dovezile constau în comportamentul lor , în ceea ce fac în interacțiunile lor cu noi (inclusiv modul în care folosesc simbolurile). Această descoperire este cu adevărat uimitoare, deoarece nimeni nu și-a imaginat că maimuțele ar putea dobândi astfel de abilități. Dar ar fi absurd să credem că atribuirea de atribute cognitive și cogitative creierului se bazează pe dovezi comparabile. Ar fi absurd pentru că nici măcar nu știm ce ar demonstra că creierul posedă astfel de atribute.

Atribuirea greșită a atributelor psihologice creierului este o formă degenerată de cartesianism.

Așadar, de ce au fost adoptate această formă de descriere și formele de explicație însoțitoare și dependente *fără argumente sau reflecții?* Bănuim că răspunsul este că a rezultat dintr -o aderență nechibzuită la o formă mutantă de cartesianism. Atribuirea de predicate psihologice minții și numai prin derivare la ființa umană a fost o caracteristică a dualismului cartezian. În reflecțiile lor asupra relației dintre descoperirile lor neurologice și capacitățile perceptive și cognitive umane , Sherrington și studenții săi, Eccles și Penfield, au rămas fideli unei forme de dualism. Succesorii lor au respins pe bună dreptate dualismul. Dar neurologii creierului de a treia generație au aplicat fără gândire creierului chiar predicatele pe care dualiștii le-au atribuit minții imateriale. Acesta a fost doar un corolar aparent inofensiv al respingerii de către neurologi a dualismului substanțial cartezian. Acești oameni de știință au continuat să explice

abilitățile perceptive și cognitive umane și exercitarea acestora pe baza exercitării, de către creier, a *propriilor abilități* cognitive și perceptive.

Atribuirea de atribute psihologice creierului este lipsită de sens.

Teza noastră este că această aplicare a predicatelor psihologice la nivelul creierului *este lipsită de sens.* Nu este vorba despre faptul că, de fapt, creierul nu gândește, nu formulează ipoteze sau nu decide, nu vede sau nu aude, nu formulează sau nu răspunde—

Nu este vorba de faptul că creierul pune întrebări, ci mai degrabă că nu are sens să atribuim astfel de predicate sau *negațiile lor* creierului. Creierul nu vede *și nici nu este orb,* așa cum bețele și pietrele nu sunt nici treze , *nici adormite.* Creierul nu aude, dar nu este surd, la fel cum nu pot fi copacii. Creierul nu ia decizii, dar nici nu este indecis. Doar ceea ce *poate* decide poate fi indecis. Astfel, nici creierul nu poate fi conștient; doar creatura din care face parte creierul poate fi conștientă sau inconștientă. *Creierul nu este un subiect logic adecvat al predicatelor psihologice.* Numai despre ființa umană și despre ceea ce *se comportă ca atare* se poate spune în mod inteligibil și literal că vede sau este orb, aude sau este surd, pune întrebări sau se abține să le pună.

Prin urmare, problema de aici este conceptuală. Nu are sens să atribuim predicate psihologice (sau negările lor) creierului, decât metaforic sau metonimic. Combinația de cuvinte rezultată nu spune nimic fals; mai degrabă, nu spune nimic, deoarece este lipsită de sens. Predicatele psihologice sunt predicate care se aplică în esență animalului viu ca întreg, nu părților sale. Nu ochiul (și cu atât mai puțin creierul) vede, ci *noi* vedem *cu* ochii noștri (și nu vedem *cu creierul* nostru , deși, fără un creier al cărui sistem vizual funcționează normal, nu am vedea). În mod similar, nu urechea aude, ci animalul căruia îi aparține urechea. Organele unui animal sunt părți ale animalului, iar predicatele psihologice sunt atribuibile întregului animal, nu părților sale constitutive.

Atribuirea de către neurocercetători a unor atribute psihologice

creierului poate fi numită „eroarea mereologică” a neuroștiinței.

Merologia este logica relației dintre parte și întreg. În neuroștiințe, vom numi eroarea neurocercetătorilor care atribuie *părților constitutive* ale unui animal atribute logic aplicabile doar animalului ca *întreg º " eroare mereologică”.* Vom numi principiul conform căruia predicatele psihologice aplicabile doar unei ființe umane (sau unui alt animal) în întregime nu pot fi aplicate inteligibil părților sale, de exemplu, creierului, un „principiu mereologic ” . Putem spune că ființele umane, dar nu și creierul lor, sunt reflexive sau nereflexive; putem spune despre animale, dar nu despre creierul lor, cu atât mai puțin despre emisferele lor, că văd, aud, miros și gust lucruri ; putem spune despre oameni, dar nu despre creierul lor, că iau decizii sau că sunt indecișI.

Este demn de remarcat faptul că multe predicate pot *fi* aplicate atât unui întreg dat (în special, unei ființe umane), cât și părților sale, iar aplicarea lor într-un caz poate fi dedusă din aplicarea lor în celălalt. Un om poate fi ars de soare, iar fața lui poate fi arsă de soare; poate tremură din tot corpul , așa că și mâinile îi vor tremura. În mod similar, uneori aplicarea unui predicat la o ființă umană nu se extinde la părțile corpului uman, de exemplu, atunci când spunem că un om a apucat clanța ușii sau, de asemenea, că mâna lui a apucat clanța; că a alunecat sau că piciorul i-a alunecat. Nu este nimic logic nepotrivit aici. Dar predicatele psihologice se aplică paradigmatic ființei *umane (sau animalului) ca întreg și nu* corpului și părților sale. Există câteva excepții, de exemplu, aplicarea verbelor care denotă senzație, cum ar fi „a răni” sau „a răni”, la părțile corpului, ca în cazurile „Mă doare mâna”, „Îmi rănești mâna”. [15] Însă varietatea predicatelor psihologice care ne interesează – adică cele pe care neurocercetătorii, psihologii și specialiștii în științe cognitive le invocă în eforturile lor de a explica capacitățile umane și exercitarea acestora – nu au o aplicație literală la părțile corpului. În special, ele nu au o aplicație inteligibilă la creier.

3.2 ÎNDOIELI METODOLOGICE

Obiecții metodologice la acuzația că neurocercetătorii comit o eroare mereologică

Dacă cineva atribuie un predicat unei entități căreia predicatul în cauză nu i se aplică logic și i se arată acest lucru , ne putem aștepta doar să insiste indignat că „nu asta am vrut să spun". La urma urmei, ar putea argumenta acea persoană , deoarece nonsensul este o combinație de cuvinte care nu spune nimic și nu descrie nicio stare de fapt posibilă, evident că nu a *vrut să spună* nonsensuri - nu se *poate* intenționa să spui nonsensuri, deoarece nu există nimic de înțeles. Așadar, cuvintele lor nu ar trebui luate în sensul lor obișnuit. Poate că expresiile problematice au fost folosite într-un sens special și sunt, de fapt, doar omonime; sau au fost *extensii analogice* ale uzului obișnuit, atât de comun în știință; sau au fost folosite metaforic sau figurativ. Dacă se poate recurge la astfel de căi de scăpare, acuzația că neurocercetătorii au căzut victimă erorii mereologice este nejustificată. Deși folosesc același vocabular psihologic ca o persoană obișnuită, îl folosesc diferit. Prin urmare, obiecțiile la adresa utilizării acestor termeni de către neurocercetători pe baza utilizării lor comune sunt irelevante.

Totuși, lucrurile nu sunt atât de simple. Este clar că persoana care atribuie greșit un predicat în acest fel nu intenționează să rostească o combinație de cuvinte fără sens. Dar faptul că nu intenționează *să rostească* prostii nu garantează că nu o vor face. Deși vor insista în mod natural că „nu au vrut să spună asta", că predicatul în cauză nu era folosit în sensul său obișnuit, insistența lor nu este cuvântul final. Autoritatea supremă în această chestiune este *propriul raționament al vorbitorului. Trebuie* să ne uităm la consecințele pe care le trag din propriile cuvinte, iar inferențele pe care le fac vor demonstra dacă foloseau predicatul într-un sens nou sau îl foloseau greșit. Dacă trebuie să fie condamnați, trebuie să fie prin propria lor recunoaștere.

Așadar, să analizăm rutele de evadare propuse, care au scopul de a demonstra că neurocercetătorii și specialiștii în științe cognitive nu sunt vinovați de erorile de care i-am acuzat.

Prima obiecție (Ullmann): predicatele psihologice astfel utilizate sunt omonime ale predicatelor psihologice obișnuite și au un sens tehnic diferit.

În primul rând, s-ar putea sublinia faptul că neurologii folosesc într-adevăr omonime, care au semnificații complet diferite. Nu este nici neobișnuit, cu atât mai puțin problematic, ca oamenii de știință să introducă un nou mod de a vorbi sub presiunea unei noi teorii. Dacă acest lucru îi derutează pe cititorii neinformați, confuzia poate fi ușor rezolvată. Evident, creierele nu gândesc, nu cred, nu deduc, nu interpretează sau nu emit ipoteze literal, ci mai degrabă gândesc*, cred*, deduc*, interpretează* sau emit ipoteze*. Ele nu au sau nu construiesc reprezentări simbolice , ci mai degrabă reprezentări simbolice*. [16]

A doua obiecție (Gregory): predicatele psihologice utilizate în acest fel sunt extensii analogice ale expresiilor obișnuite

În al doilea rând, s-ar putea sublinia faptul că neurologii extind utilizarea comună a vocabularului relevant prin analogie, așa cum s-a făcut adesea de-a lungul istoriei științei, de exemplu, în extinderea analogică a hidrodinamicii la teoria electricității. Prin urmare, obiecția față de atribuirea predicatelor psihologice creierului pe motiv că în vorbirea cotidiană astfel de predicate sunt aplicabile doar animalului în ansamblu ar echivala cu demonstrarea unei forme de *inerție semantică?*

A treia obiecție (Blakemore): atribuirea pe care neurologii o fac creierului a atributelor psihologice este figurativă sau metaforică, deoarece ei știu perfect că acesta nu gândește și nu folosește hărți.

În cele din urmă, s-ar putea argumenta că neurocercetătorii nu cred, *de fapt,* că creierul

raționează, argumentează, formulează și răspunde la întrebări în același mod în care o facem noi. Ei nu cred cu adevărat că creierul interpretează indicii, face conjecturi sau conține simboluri care descriu lumea exterioară. Și, deși vorbesc despre existența „hărților" în creier și că acesta conține „ reprezentări interne", ei nu folosesc aceste cuvinte în sensul lor obișnuit sau comun. Este limbaj figurativ și metaforic, uneori chiar licență poetică. [18] Prin urmare, neurocercetătorii nu sunt deloc derutați de un astfel de limbaj; știu perfect ce vor să spună, dar le lipsesc cuvintele pentru a o spune, decât dacă o fac metaforic sau figurativ.

Răspuns la obiecția conform căreia neurologii folosesc vocabularul psihologic într-un sens tehnic special

În ceea ce privește utilizarea greșită a vocabularului psihologic implicită în atribuirea predicatelor psihologice creierului, toate dovezile indică faptul că neurocercetătorii nu folosesc acești termeni într-un sens special. Expresiile psihologice pe care le folosesc, departe de a fi omonime noi, sunt invocate în sensul lor obișnuit; altfel, neurocercetătorii nu ar deduce din ele ceea ce deduc. Când Crick afirmă că „ceea ce vedem nu este *cu adevărat* ceea ce este acolo; este ceea ce creierul nostru *crede* că este acolo", este important ca el să presupună că „crede" are conotațiile sale normale - că nu înseamnă același lucru ca un anumit termen nou, „crede*"'. Căci face parte din teoria lui Crick că credința este rezultatul unei *interpretări* bazate pe *experiențe și* informații anterioare (și nu rezultatul unei interpretări* bazate pe experiențe* și *informații anterioare* *). Când Semir Zeki subliniază faptul că dobândirea de cunoștințe este o „funcție primară a creierului", [el] se referă la cunoaștere (nu la cunoaștere*), altfel nu ar crede că este sarcina neuroștiinței viitorului să rezolve problemele epistemologiei (ci doar, probabil, ale epistemologiei *). În mod similar, când Young afirmă că creierul conține cunoștințe și informații, care sunt codificate în creier „în același mod în care cunoștințele sunt înregistrate în cărți sau computere", el se referă la cunoaștere, nu la cunoaștere* - deoarece cunoștințele [și] informațiile, nu cunoștințele* și informațiile*, pot fi înregistrate în cărți și computere. Când Milner, Squire și Kandel vorbesc despre „memorie

declarativă", ei clarifică faptul că această expresie înseamnă „ceea ce se înțelege în mod normal prin termenul «memorie»", [dar] apoi afirmă că amintirile asociate cu o astfel de memorie (nu cu memoria*) sunt „stocate în creier". Aceasta presupune că are sens să vorbim despre memorie (în sensul obișnuit al cuvântului) ca despre activitatea de stocare a amintirilor *în creier.* [11]

Un răspuns la Ullmann: David Marr despre „reprezentări"

Acuzația de comitere a erorii mereologice nu poate fi respinsă atât de ușor. Dar s-ar putea părea că Simon Ullman se bazează pe un teren mai solid atunci când vorbește despre reprezentări interne și reprezentări simbolice (precum și hărți) *în creier.* Dacă „ -reprezentare" nu înseamnă ceea ce înseamnă în mod normal, dacă „simbolic " nu are nicio legătură cu simbolurile, atunci, într-adevăr, ar putea fi inofensiv să spunem că există -reprezentări interne și simbolice în creier. (Și dacă „hărțile" nu au nicio legătură cu atlasele, ci doar cu *operațiunile de cartografiere,* (Atunci poate fi, de asemenea, inofensiv să spunem că există hărți în creier.) Multiplicarea omonimelor este extrem de nerecomandată, dar nu implică neapărat incoerență conceptuală, *atâta timp cât oamenii de știință care folosesc acești termeni în acest fel nu uită că termenii nu au sensul lor obișnuit.* Din păcate, este obișnuit ca aceștia să uite acest lucru și *să procedeze la confundarea noii utilizări cu cea veche,* căzând astfel în incoerență. Ullman, în apărarea lui Marr, insistă (perfect corect) că anumite evenimente cerebrale pot fi considerate reprezentări* ale adâncimii , orientării sau reflectanței, [23] adică, că anumite modele de activare neuronală pot fi corelate cu caracteristici ale câmpului vizual (și vorbesc despre primele ca „reprezentări*" ale celui de-al doilea). Dar este evident că nu asta este tot ce vrea să spună Marr. El susține că sistemele numerice (cifre romane sau arabe, notația binară) sunt reprezentări. Cu toate acestea, astfel de notații nu au nicio legătură cu corelațiile cauzale, ci cu convențiile reprezentaționale. El susține că „o reprezentare a formei ar fi o schemă formală descriptivă a unor aspecte ale formei, împreună cu reguli care specifică modul de aplicare a schemei la o anumită formă"; [24] că o schemă formală

este „un set de simboluri și regulile pentru combinarea lor”; [25] și că „prin urmare, o reprezentare nu este deloc o idee ciudată ; cu toții folosim reprezentări în mod continuu. Cu toate acestea , noțiunea că cineva poate înțelege un anumit aspect al realității făcând o descriere a acestuia și folosind un simbol în acest scop și că acest lucru poate fi util, mi se pare o idee puternică și fascinantă ” .[26] Dar sensul în care „folosim reprezentări în mod continuu”, *în care reprezentările sunt simboluri guvernate de reguli și în care sunt folosite pentru a descrie lucruri,* este *sensul semantic al „reprezentării”, nu un nou sens omonim bazat pe o corelație cauzală.* [Marr] cade în propria capcană.27 În realitate , el confundă - reprezentările* lui Ullman , care *sunt* corelații cauzale, cu reprezentările, care sunt simboluri sau sisteme de simboluri cu o sintaxă și un sens determinate prin convenție.

Răspuns către Ullman: Young despre „hărți” și Frisby despre „reprezentări simbolice”

În mod similar, ar fi o eroare inofensivă să vorbim despre hărți cerebrale atunci când ceea ce se înțelege este că anumite caracteristici ale câmpului vizual pot fi mapate pe descărcările grupurilor de celule din cortexul vizual primar sau striat. Dar din aceasta nu se poate deduce, așa cum face Young, că creierul își folosește hărțile atunci când formulează ipoteze despre ceea ce este vizibil. De asemenea, ar fi inofensiv să spunem că există reprezentări simbolice în creier, cu condiția ca „simbolic” să nu aibă nicio legătură cu sensul semantic, ci să însemne pur și simplu „sens natural” (ca în „fumul înseamnă foc”). Dar atunci nu putem adăuga, așa cum face Frisby, că „în creier trebuie să existe o descriere simbolică a lumii exterioare, o descriere formulată în simboluri care se referă la diversele aspecte ale lumii de care văzul ne face conștienți”. [28] Căci această utilizare a „simbolului” este în mod clar semantică. Deoarece, chiar dacă fumul semnifică foc, în sensul că este un semn *de* foc (o indicație corelată cu focul prin inducție), nu este un semn *de* foc în sine. Fumul care se ridică de pe un deal îndepărtat nu este o descriere a focului formulată în simboluri, nici activarea neuronilor din cortexul striat „vizual” nu este o descriere simbolică a obiectelor din câmpul vizual, chiar dacă un neurocercetător poate face inferențe empirice despre ceea ce este

vizibil pentru un animal pe baza cunoștințelor sale despre celulele activate în cortexul său striat „vizual". Activarea celulelor din VI poate fi un semn al unei figuri cu anumite orientări liniare în câmpul vizual al animalului, dar nu *semnifică* nimic, nu este un simbol și nu descrie nimic.

Un răspuns la a doua obiecție este că, atribuind atribute psihologice creierului, neurologii nu comit eroarea mereologică, ci pur și simplu extind vocabularul psihologic analogic.

O altă modalitate de a respinge acuzația că descrierile neurocercetătorilor asupra descoperirilor lor depășesc adesea limitele sensului este de a argumenta că utilizarea neuroștiințifică, departe de a fi incoerentă din punct de vedere conceptual, este inovatoare, deoarece extinde vocabularul psihologic în moduri noi. Este cu siguranță adevărat că analogiile sunt o sursă de înțelegere științifică. Analogia hidrodinamică s-a dovedit utilă în dezvoltarea teoriei electrice, chiar dacă curentul electric nu curge în același mod ca apa, iar un fir electric nu este o țeavă de niciun fel. Ceea ce este discutabil este dacă aplicarea vocabularului psihologic la creier ar trebui înțeleasă ca analogică.

Perspectivele ca acest lucru să nu se întâmple nu par bune. Aplicarea expresiilor psihologice la creier nu face parte dintr-o teorie complexă, plină de relații funcționale și matematice care pot fi exprimate prin legi cuantificabile, precum cele care trebuie să existe în teoria electricității. Se pare că este nevoie de ceva mult mai flexibil. Prin urmare, este adevărat că psihologii, urmându-l pe Freud și pe alții, au extins conceptele de credință, dorință și motiv pentru a vorbi despre *credințe, dorințe și motive inconștiente.* Atunci când aceste concepte sunt supuse unei astfel de extinderi, apare ceva nou care necesită explicații. Expresiile nou extinse nu mai admit aceleași posibilități combinatorii ca înainte. Ele au un sens diferit, care are o relație importantă cu cel anterior și care necesită explicații. Relația dintre o credință (conștientă) și o credință inconștientă, de exemplu, nu este aceeași cu relația dintre un scaun vizibil și

> un scaun ascuns; nu este „pur și simplu ca o credință conștientă, doar inconștientă", ci mai degrabă ca relația dintre „Vi" și „Vl". Însă atunci când neurocercetătorii precum Sperry și Gazzaniga spun că emisfera stângă ia decizii, generează interpretări, cunoaște, observă și explică lucrurile , devine clar din ceea ce adaugă în continuare că acestor expresii psihologice nu li s-a dat un nou sens. Altfel, nu ar spune ei că o emisferă cerebrală este „un sistem conștient în sine, care percepe, gândește, își amintește, rationează , dorește și simte, *toate la un nivel caracteristic uman"?* [3]

Nu inerția semantică ne motivează afirmația că neurocercetătorii se angajează în diverse forme de inconsecvență . Mai degrabă, este recunoașterea cerințelor logicii asupra expresiilor psihologice. Predicatele psihologice pot fi predicate doar despre un animal întreg, nu despre părțile sale. Nu au fost stabilite convenții pentru a determina ce se poate înțelege prin atribuirea unor astfel de predicate unei părți a unui animal, în special creierului său. Astfel, aplicarea unor astfel de predicate creierului sau emisferelor sale depășește limitele sensului. Afirmațiile rezultate nu sunt false, deoarece pentru a spune că ceva este fals trebuie să avem o idee despre ce ar implica dacă ar fi adevărat - în acest caz, ar trebui să știm ce ar implica dacă creierul ar gândi , ar raționa, ar vedea și ar auzi etc. și am descoperit că, de fapt, creierul nu face astfel de lucruri. Dar nu avem această idee, iar aceste afirmații nu sunt false. Mai degrabă, frazele în cauză sunt lipsite de sens. Aceasta nu înseamnă că sunt ridicole sau prostești. Înseamnă că acestor combinații de cuvinte nu li s-a atribuit nicio semnificație ***și,*** prin urmare, nu spun absolut nimic, chiar dacă ar putea părea că spun asta.

Răspuns la a treia obiecție (Blakemore) conform căreia aplicarea predicatelor psihologice la creier este doar metaforică

A treia obiecție metodologică a fost ridicată de Colin Blakemore. Referitor la observația lui Wittgenstein conform căreia „numai despre ființele umane ***și*** despre ceea ce seamănă cu o ființă umană (se comportă ca atare) se poate spune: are senzații; vede, este oarbă; aude, este surdă; este conștientă sau inconștientă ", Blakemore subliniază că „este

trivială, poate pur și simplu eronată". În ceea ce privește acuzația că afirmația neurocercetătorilor conform căreia există „hărți" în creier este plină de potențial de - confuzie (deoarece tot ce se poate spune este că este posibil să se cartografieze , de exemplu, aspecte ale elementelor din câmpul vizual pe activarea celulelor din cortexul striat „vizual"), Blakemore observă că există dovezi copleșitoare ale „modelelor topografice de activitate" în creier.

> Încă din vremea lui Hughlings Jackson, conceptul de subdiviziune funcțională și reprezentare topografică a devenit indispensabil pentru cercetarea creierului. Sarcina de cartografiere a creierului este departe de a fi completă, dar succesele din trecut ne dau motive să credem că fiecare parte a creierului (și în special cortexul cerebral) *este probabil* organizată topografic. La fel ca în cazul descifrării unui cod, traducerii sistemului de scriere Linear B sau citirii hieroglifelor, tot ce ne trebuie pentru a recunoaște ordinea creierului este un set de reguli - reguli care leagă activitatea nervoasă de evenimentele care au loc în lumea exterioară sau în corpul animalului. [30]

Termenul „reprezentare", apropo, aici înseamnă pur și simplu conectivitate cauzală sistematică. Este inofensiv. Dar nu trebuie confundat cu sensul în care se poate spune că o propoziție dintr-o limbă reprezintă starea de lucruri pe care o descrie, că o hartă reprezintă ceea ce este o hartă sau că o pictură reprezintă ceea ce este o pictură. Cu toate acestea, o astfel de ambiguitate în utilizarea cuvântului „reprezentare" este periculoasă, deoarece este probabil să ducă la confuzie între diferitele sensuri. Următoarele observații ale lui Blakemore demonstrează cât de confuz poate fi acest lucru:

> Având în vedere aceste dovezi copleșitoare ale tiparelor topografice ale activității cerebrale, nu este surprinzător faptul că neurofiziologii și neuroanatomiștii au mers atât de departe încât să afirme că există hărți *care se presupune că* joacă un rol esențial în reprezentarea și interpretarea lumii de către creier, la fel cum o fac hărțile dintr-un atlas pentru cititorii săi. Biologul J.Z. Young descrie creierul ca posedând un limbaj pictografic: „Ceea ce se întâmplă în creier trebuie să ofere o reprezentare fidelă a

evenimentelor care au loc în afara acestuia, iar aranjamentul celulelor din interiorul său oferă un model detaliat al lumii. Creierul comunică semnificații prin analogii topografice ." ³¹ Există însă vreun pericol în utilizarea metaforică a unor termeni precum „limbaj", „gramatică" și „hartă" pentru a descrie proprietățile creierului?... Nu-mi pot imagina vreun neurofiziolog crezând că există un cartograf fantomă care răsfoiește un atlas al creierului . Nici nu cred că utilizarea cuvintelor din limbajul comun (cum ar fi hartă, reprezentare, cod, informație și chiar limbaj) este o eroare conceptuală gravă de acest [presupus] fel. O astfel de imagistică metaforică este un amestec de descriere empirică, licență poetică și vocabular inadecvat. [32]

Existența oricărui pericol potențial în utilizarea metaforică a cuvintelor depinde de cât de clar este că este doar metaforică și de intenția autorului ca aceasta să *fie* tot ce există. Este foarte îndoielnic că atributele pe care neurocercetătorii le atribuie creierului - atribute care, la propriu, pot fi aplicate doar unui animal în ansamblu - sunt cu adevărat doar metaforice (metonimie sau sinecdocă). Neurofiziologii, *desigur, nu cred* într-un „cartograf fantomă" care studiază cu atenție un atlas cerebral, dar cred că creierul *folosește* hărți. Potrivit lui Young, creierul *construiește ipoteze și* o face *pe baza* acestei „reprezentări organizate topografic". [33] Întrebarea este: ce concluzii trag neurocercetătorii din afirmația lor că există hărți sau reprezentări în creier sau din teza lor că creierul conține informații sau din ideea că există (așa cum sugerează J.Z. Young) „limbaje ale creierului"? Aceste utilizări presupus metaforice sunt în mare măsură capcane pentru utilizator. Utilizatorul *nu va cădea neapărat* în ele, dar este foarte probabil să o facă.

Confuzia lui Blakemore

Citatul precedent din Blakemore subliniază cât de ușor poate duce la confuzie ceea ce se presupune a fi o metaforă inofensivă. Pentru că, deși poate fi inofensiv să vorbim despre „hărți", adică despre *cartografiere,* Având în vedere că relațiile topografice dintre

grupurile de celule care sunt sistematic legate de caracteristicile câmpului perceptiv joacă un rol în *interpretarea* acestor caracteristici de către creier, nu este inofensiv să spunem că astfel de „hărți" joacă „un rol esențial în *reprezentarea* și *interpretarea* lumii de către creier, la fel cum o fac hărțile dintr-un atlas pentru cititorii săi" (sublinierea noastră). În primul rând, nu este clar ce semnificație are termenul „interpretare" în acest context. Nu este nicidecum evident ce se înțelege prin afirmația că relațiile topografice dintre grupurile de celule care sunt sistematic legate de caracteristicile câmpului perceptiv joacă un rol esențial în *interpretarea* de către creier a ceva. A interpreta, în sensul său literal, înseamnă a explica sensul a ceva sau a considera că ceva ambiguu are un sens și nu altul. Dar nu are sens să presupunem că creierul *explică* ceva sau că înțelege că ceva *înseamnă* un lucru și nu altul. Dacă ne uităm la JZ Young pentru a afla ce avea în minte, ceea ce descoperim este afirmația că creierul „construiește ipoteze și programe" din astfel de hărți, ceea ce nu face decât să ne afunde și mai adânc în mlaștină.

Și, mai important, indiferent de sensul pe care l-am da afirmației lui Blakemore conform căreia „hărțile creierului" (care nu sunt de fapt hărți) joacă un rol esențial în „reprezentarea și interpretarea" lumii de către creier, *aceasta nu poate* fi *„în același mod în care o fac hărțile dintr-un atlas pentru cititorii lor".* Căci o hartă este o reprezentare picturală , realizată conform convențiilor de cartografiere și regulilor de proiecție. Oricine poate citi un atlas trebuie să cunoască și să înțeleagă aceste convenții și să extragă, din hărți, - caracteristicile a ceea ce este reprezentat. Dar „hărțile" creierului nu sunt deloc hărți în acest sens. Creierul nu este același lucru cu un cititor de hărți, deoarece nu se poate spune că cunoaște vreo convenție despre reprezentări sau metode de proiecție și nici că este capabil să citească ceva din aranjamentul topografic al celulelor într-o stare activată și conform unui set de convenții. Celulele nu sunt deloc aranjate conform convențiilor, iar corelația dintre starea lor de activare și caracteristicile câmpului perceptiv nu este convențională, ci *cauzală.*

FUNDAMENTE FILOZOFICE DE NEUROȘTIINȚE

Extras din Capitolul 10

MAXWELL BENNETT și PETER HACKER

10.3 *QUALIA*

Qualia înțeleasă *ca caracter calitativ al experienței: concepția filosofilor*

Tentația de a extinde conceptul de conștiință pentru a cuprinde întregul domeniu al „experienței" a fost întărită de introducerea eronată de către filosofi a ideii de *qualia.* Din păcate , neurocercetătorii au adoptat această idee aberantă și neînțelegerile asociate cu ea. Termenul „ *qualia"* a fost introdus pentru a se referi la presupusul „caracteru calitativ al experienței ". Se spune că fiecare experiență are un caracter calitativ distinct . *Qualia,* susține Ned Block, „includ ceea ce se simte când vezi, auzi și miroși, cum se simte când simți durere; mai general, cum este să ai stări mentale. *Qualia* sunt proprietăți experiențiale ale senzațiilor, sentimentelor, percepțiilor și [...], de asemenea, ale gândurilor și dorințelor." [1] În mod similar, Searle susține că „fiecare stare conștientă poartă cu sine un anumit sentiment calitativ , iar acest lucru poate fi observat luând în considerare câteva exemple. Experiența degustării berii este foarte diferită de cea a ascultării"

Simfonia a IX-a de Beethoven și ambele au un caracter *calitativ* distinct față de mirosul unui trandafir sau contemplarea unui apus de soare. Aceste exemple ilustrează diferitele caracteristici calitative ale experiențelor conștiente." [2] La fel ca Block, Searle susține, de asemenea, că gândirea poartă cu sine o senzație calitativă specială. „Există ceva asemănător cu gândirea că doi plus doi egal patru . Nu există nicio modalitate de a o descrie decât spunând că este caracterul gândirii conștiente că «doi plus doi egal patru»." [3] „Cel mai bun mod de a caracteriza" obiectul unei investigații a conștiinței, subliniază Chalmers, este „calitatea subiectivă a experienței ". O stare mentală, spune el, este conștientă „dacă are o *senzație calitativă,* o calitate a experienței asociată cu ea. Aceste senzații calitative sunt cunoscute și sub numele de calități fenomenale sau, mai pe scurt, *qualia. „* Problema explicării acestor calități fenomenale nu este nimic mai mult decât problema explicării conștiinței." [4] De asemenea, el susține opinia că gândirea este o experiență cu conținut calitativ: „Când mă gândesc la un leu, de exemplu, fenomenologia mea pare să aibă o calitate leonină: ceea ce simte cineva când se gândește la un leu este subtil diferit de ceea ce simte cineva când se gândește la Turnul Eiffel." [5]

Neurologii *îi urmează pe filosofi*

Neurologii au susținut ideea de *qualia.* Ian Glynn susține că „deși *qualia sunt* cel mai clar asociate cu senzațiile și percepțiile, ele se găsesc și în alte stări mentale, cum ar fi credințele, dorințele, speranțele și temerile, în timpul episoadelor conștiente ale acestor stări." [6] Damasio afirmă că „ *qualia* sunt calitățile senzoriale simple pe care le găsim în albastrul cerului sau în tonalitatea unui violoncel și, în consecință, componentele fundamentale ale imaginilor [în care se presupune că constă percepția] constau din *qualia."* Edelman și Tononi susțin că „fiecare experiență conștientă diferențiabilă reprezintă un *quale"* diferite, fie că este vorba în primul rând de o senzație, o imagine, un gând sau chiar o dispoziție [...]", [8] și apoi afirmând că „problema *qualia"* este „poate cea mai descurajantă problemă a conștiinței ".

Explicația naturii calitative a experienței din perspectiva faptului că trăirea acesteia face ca cineva să simtă ceva

Senzația subiectivă sau calitativă a unei experiențe conștiente este caracterizată, la rândul ei, din perspectiva că *ceva este simțit* atunci când un organism are o experiență. „O experiență sau altă entitate mentală subiectivă este «conștientă fenomenal»", spune *Enciclopedia Routledge de Filosofie* , „ numai dacă există ceva ce cineva simte având -o" [.9] „Stările conștiente sunt calitative", explică Searle, „în sensul că pentru fiecare stare conștientă [...] există ceva care este simțit calitativ în a fi în acea stare".10 [Ideea] și utilizarea mistificatoare a expresiei „există ceva care este simțit" sau „există ceva care este ca..." provin dintr-un articol al filosofului Thomas Nagel intitulat „Cum este să fii liliac?". Nagel susține că „faptul că un organism are experiențe conștiente înseamnă practic că există ceva de simțit în legătură cu faptul de *a fi* acel organism [...] Fundamental, un organism are stări mentale conștiente dacă și numai dacă există ceva de simțit în legătură cu faptul de *a fi* acel organism, dacă a fi acel organism este ceva *pentru* organism". [11] Aceasta, adică ceea ce simte organismul în legătură cu faptul că este acel organism, este caracterul subiectiv sau calitatea experienței.

Explicația lui Nagel asupra conștiinței din perspectiva că „există ceva care se simte ca..."

Dacă presupunem că înțelegem sintagma „există ceva care se simte ca" folosită în acest fel, atunci se pare că ideea lui Nagel ne clarifică conceptul de creatură conștientă și pe cel de experiență conștientă:

1) *O creatură este conștientă sau are experiență dacă și numai dacă este în scopul reproducerii*

Este ceva la ea care da senzația că este creatura care este.

2) *O experiență este o experiență conștientă dacă și numai dacă, pentru subiect*

Pentru cei care o au, există ceva ce seamănă cu experiența.

Așadar, pentru un liliac, există ceva care se simte ca și cum ai fi liliacul (deși Nagel susține că nu ne putem imagina *ce* este) și există ceva care se simte ca și cum ai fi o ființă umană (și susține că știm cu toții cum este să fim noi înșine).

Este important de menționat că sintagma „există ceva *care se simte ca și cum* ai avea o experiență E pentru un subiect" *nu* indică *o comparație.* Nagel nu spune că a avea o anumită experiență conștientă *este similar* cu ceva (de exemplu, o altă experiență), ci mai degrabă că există ceva care se simte ca și cum ai avea-o *pentru subiect;* adică, „care se simte ca..." înseamnă „cum se simte pentru subiect". [12] Este surprinzător, însă, că Nagel nu spune niciodată, nici măcar în legătură cu o singură experiență, cum este să o ai. El susține că caracterul calitativ al experiențelor altor specii poate fi dincolo de capacitatea noastră de a concepe. De fapt, același lucru se poate spune și despre experiențele altor ființe umane. „Caracterul subiectiv al experienței unei persoane surde și oarbe din naștere, de exemplu, nu este accesibil mie și nici a mea nu este accesibilă lor". Dar știm cum este să fii noi „și, deși nu posedăm vocabularul pentru a o descrie în mod adecvat, caracterul său subiectiv este extrem de specific și, în anumite privințe, descriptibil în termeni pe care doar creaturile ca noi îi pot înțelege". [13]

Filosofii și *neurologii sunt de acord*

Filosofii și neurocercetătorii au susținut această idee. Ei consideră că surprinde natura esențială a ființelor conștiente și a experienței conștiente. Astfel, Davies și Humphries susțin că „deși nu există nimic care să se simtă ca o cărămidă sau ca o imprimantă cu jet de cerneală, probabil există ceva care să se simtă ca un liliac sau un delfin și cu siguranță există ceva care să se simtă ca o ființă umană. Un sistem, fie el o creatură sau un artefact, este conștient numai dacă există ceva care să se simtă ca fiind acel sistem" .[14] Edelman și Tononi sunt de acord că „știm deja cum este să fim noi înșine, dar vrem să explicăm de ce

suntem conștienți, de ce există «ceva» care este ca și cum ai fi noi înșine, să explicăm, pe scurt, cum sunt generate calitățile experiențiale subiective". [15] Iar Glynn susține că, în ceea ce privește experiențele noastre, de exemplu, mirosul cafelei proaspăt măcinate, ascultarea notelor oboiului sau contemplarea albastrului cerului, „știm cum este să ai aceste experiențe doar pentru că le avem sau le-am avut. [...] Așa cum simțim într-un anumit fel când mirosim cafea proaspăt măcinată, putem simți într-un anumit fel (cel puțin intermitent) și atunci când credem că... sau dorim că... sau ne temem că...".

Astfel, *qualia* sunt concepute ca fiind caracteristicile calitative ale „stărilor mentale" sau „experiențelor", categorii care includ nu doar percepția, senzația și afectul, ci și - dorința, gândul și credința. Pentru fiecare „experiență conștientă" sau fiecare „stare mentală conștientă", există ceva ce se simte ca și cum ai avea-o sau ai fi în ea. Acest ceva este un *qualia.* un „sentiment calitativ". „Problema explicării acestor calități fenomenale ", spune Chalmers, „nu este alta decât problema explicării conștiinței". [16]

10.31 „CUM E" SĂ AI O EXPERIENȚĂ

Principalele motive pentru extinderea conceptului comun de conștiință

Unul dintre motivele oferite pentru extinderea conceptului de conștiință dincolo de limitele sale prudente și legitime este că ceea ce este distinctiv, remarcabil și, fără îndoială, misterios în legătură cu experiențele este *faptul că există ceva anume în a le avea.* Se susține că o experiență *este o* experiență *conștientă* numai dacă, pentru subiectul experienței, există ceva care „se simte ca" să o aibă. Conștiința , astfel concepută, este *definită* în termenii *sentimentului calitativ al experienței.* Există un mod specific de a simți atunci când vezi, auzi și miroși, când simți durere sau chiar când „ai stări mentale" (Block); fiecare stare conștientă implică un anumit *sentiment calitativ* (Searle), iar fiecare experiență conștientă distinctă reprezintă o *calitate.* distinctă (Edelman și Tononi).

Această senzație calitativă, unică oricărei experiențe distincte, este ceea ce *simte - subiectul experienței având experiența. Sau cel puțin* , acesta este argumentul.

Folosirea unor astfel de expresii ciudate pentru a invoca ceva ce se presupune că ne este familiar tuturor ar trebui să ne ridice suspiciuni. Mai întâi , vom analiza „modurile de a simți" și apoi acel „ceva care se simte ca..."

Există întotdeauna o modalitate specifică de a simți că ai o „experiență conștientă"?

Există într-adevăr un *mod specific* în care se simte văzul, auzul sau mirosul? De fapt, putem întreba pe cineva care și-a recăpătat vederea, auzul sau simțul mirosului: „Cum este să vezi (să auzi, să miroși) din nou?" Ne-am putea aștepta ca persoana să răspundă: „Este minunat" sau poate: „Este o senzație foarte ciudată". Întrebarea se referă la atitudinea persoanei față de experiența sa și la simțul său de percepție recăpătat - de exemplu, că pare fantastic să poți vedea din nou sau ciudat să auzi din nou după atâția ani de surditate. În aceste cazuri, există un mod în care se simte văzul sau auzirea din nou; mai exact, este fantastic sau ciudat. Dar dacă am întreba o persoană obișnuită cum se simte să vadă masa, scaunul, biroul, covorul etc., s-ar întreba ce încercăm să obținem. Nu este nimic distinctiv în a vedea aceste obiecte de zi cu zi. Evident , a vedea o masă diferă de a vedea un scaun, un birou, un covor etc., dar diferența nu constă în faptul că a vedea un birou *se simte diferit* față de a vedea un scaun. A vedea o masă sau un scaun obișnuit nu provoacă *nicio* reacție emoțională sau atitudinală în circumstanțe normale. Experiențele diferă în măsura în care obiectele lor diferă.

Am putea spune, stângaci, că există o modalitate de a simți durerea. Aceasta este doar o modalitate complicată de a spune că există un răspuns la întrebarea (destul de prostească) „Cum este să ai durere?" - de exemplu, că este foarte neplăcută sau, în unele cazuri, teribilă. Așadar, s-ar putea spune că atunci când suferi de migrenă, simți ceva distinctiv, mai exact ceva foarte neplăcut . Această afirmație este inofensivă, dar nu oferă nicio dovadă pentru teza generală conform căreia, pentru fiecare experiență distinctă, există o modalitate specifică de a o simți. Durerile sunt o excepție, deoarece, prin definiție, au o tonalitate hedonic negativă. Durerile sunt senzații intrinsec neplăcute. Cu toate acestea, a percepe nu înseamnă a avea senzații. Iar percepția, în diversele sale forme și cu numeroasele sale obiecte posibile, nu este în mod normal, deși poate fi ocazional, subiectul vreunei calități afective sau atitudinale (de exemplu, plăcut, agreabil, oribil). Nici subiectul nu are o calitate diferită pentru fiecare obiect al fiecărei modalități perceptive. Mai mult, pentru o mare diversitate de lucruri care pot fi numite „experiențe", nu există un „mod specific de a simți " că le ai; adică, nu există un răspuns la întrebarea „Cum te simți să...?"

Nu putem să nu fim de acord cu Searle că experiența degustării berii este destul de diferită de cea a ascultării *Simfoniei a Noua* de Beethoven și că ambele sunt diferite de mirosul unui trandafir sau de contemplarea unui apus de soare. Experiențele perceptive sunt în esență identificate sau specificate prin modalitatea lor - văzul, auzul, gustul, mirosul și atingerea - și prin obiectele lor - ceea ce le reprezintă. Însă a susține că diversele experiențe au o *senzație unică și distinctivă* este complet diferit și mult mai discutabil. Este mai discutabil în măsura în care nu este clar ce *înseamnă.* Este evident că cele patru experiențe pe care le citează Searle sunt în mod normal plăcute pentru mulți oameni. Și este întru totul adevărat că identitatea plăcerii sau a încântării depinde de obiectul plăcerii. Nu se poate obține plăcerea de a bea bere ascultând *Simfonia a Noua de Beethoven* și nici plăcerea de a contempla un apus de soare din parfumul unui trandafir. Acesta este un adevăr logic, nu empiric; Adică, nu este vorba despre faptul că senzația calitativă distinctă a contemplării unui apus de soare diferă factual de senzația distinctă a mirosului unui trandafir - la urma urmei, ambele pot fi foarte plăcute. Ceea ce se întâmplă este că plăcerea de a vedea un apus de soare diferă logic de cea a mirosului

unui trandafir, deoarece identitatea plăcerii depinde de ceea ce o produce . Nu rezultă din aceasta că fiecare experiență are un *caracter calitativ diferit* , adică că există o senzație specifică pentru toate experiențele. Deoarece, în primul rând, majoritatea experiențelor nu au, în acest sens, niciun caracter calitativ; nu produc nici plăcere, nici neplăcere, nu sunt nici plăcute, nici neplăcute etc. Pe măsură ce mergem pe stradă, vedem sute de obiecte diferite. A vedea un felinar este o experiență diferită de a vedea o cutie poștală, dar simțim ceva diferit în fiecare caz? Nu; Nici nu simt la fel, deoarece vederea ambelor obiecte nu provoacă niciun răspuns ; nu există niciun sentiment calitativ asociat cu vederea niciunuia dintre ele. În al doilea rând, experiențe diferite care *au* o senzație calitativă asociată, de exemplu, cu o componentă hedonică, pot avea aceeași senzație asociată. Ceea ce le diferențiază nu este modul în care se simt, chiar dacă întrebarea „Cum te simți să-l vezi pe V?" (unde „V" specifică o experiență adecvată) poate avea exact același răspuns, deoarece experiențe diferite pot fi la fel de plăcute sau neplăcute, interesante sau plictisitoare.

Natura calitativă a experiențelor concepute corect

Atât simțirea durerii, cât și percepția a tot ceea ce percepe cineva pot fi numite „experiențe". A fi într-o anumită stare emoțională poate fi, de asemenea, numit o experiență. Și, bineînțeles , efectuarea unei varietăți nelimitate de activități poate fi, de asemenea, numită experiențe. Am putea spune că experiențele sunt subiectele posibile ale predicatelor atitudinale; adică, pot fi plăcute sau neplăcute, interesante sau plictisitoare, minunate sau teribile. Aceste atribute pot fi numite „caracteristici calitative ale experiențelor", nu experiențele în sine. Prin urmare, nu se poate spune în mod inteligibil că a vedea culoarea roșie sau a contempla *Guernica,* a auzi un sunet sau a asculta *Tosca* sunt *„qualia".* Prin urmare, atunci când Damasio spune că caracterul albastru al cerului este un *quale,* el schimbă sensul termenului *„quale",* deoarece dacă culoarea unui obiect este un *quale,* Astfel *, qualia nu* sunt nicidecum caracteristicile calitative ale experiențelor, ci mai degrabă calitățile obiectelor experienței (sau, dacă culorile nu sunt considerate calități ale obiectelor, atunci *quali* sunt constituenții conținutului experiențelor perceptive). În mod similar, atunci când Edelman și Tononi

afirmă că fiecare experiență conștientă distinctă reprezintă un *quale distinct* , fie că este o senzație, o imagine, o dispoziție sau un gând, ei schimbă sensul termenului *„quale”*. Căci este evident că *„quale”* nu înseamnă „caracterul calitativ al unei experiențe” în sensul pe care l-am examinat. Ceea ce înseamnă sau ar trebui să însemne este ceva ce vom analiza în scurt timp (§ 10.34).

Este demn de remarcat faptul că a spune că o experiență este subiectul unui predicat atitudinal este un mod de a vorbi potențial înșelător. Deoarece a spune că o experiență (de exemplu, a vedea, a observa, a zări, a auzi, a gusta cutare sau cutare lucru, dar și a merge, a conversa, a dansa, a juca, a urca munți, a lupta în bătălii, a picta tablouri) este însoțită de un anumit sentiment (de exemplu , că este plăcută, încântătoare, fermecătoare, neplăcută, respingătoare , dezgustătoare) înseamnă pur și simplu a spune că subiectul experienței - adică persoana care a văzut, a auzit, a gustat, a mers, a conversat, a dansat etc. - a găsit-o plăcută, încântătoare, fermecătoare etc. Prin urmare, caracterul calitativ al unei experiențe E - adică ceea ce simte cineva atunci când are acea experiență - este atitudinea afectivă a subiectului (cum s-a simțit) atunci când experimentează E.

Pentru a evita confuziile în acest moment, trebuie să distingem patru aspecte:

1) Multe experiențe sunt în esență individualizate, adică sunt distinse, specificând ceea ce reprezintă o experiență.

2) Fiecare experiență este un subiect *potențial* al unor predicate atitudinale pozitive și negative , de exemplu, predicate de plăcere, interes sau atracție. Nu rezultă din aceasta și este, de asemenea, fals că fiecare experiență este un subiect *real* al unui predicat atitudinal pozitiv sau negativ .

3) Este posibil ca experiențe diferite, fiecare supusă unui atribut atitudinal, să nu poată fi distinse prin modul în care se simte persoana care le experimentează. Mirosul trandafirilor este diferit de cel al liliacului. A mirosi trandafirii și a mirosi liliacul sunt experiențe diferite . Nu se poate obține plăcerea de a mirosi un trandafir mirosind un liliac. Dar este foarte posibil ca experiențele să fie la fel de plăcute. Așadar, dacă suntem întrebați cum se simte mirosind un trandafir și cum se simte mirosind un liliac, răspunsul ar putea fi perfect același,

și anume „ceva încântător”. Dacă acest răspuns specifică care a fost senzația, atunci este evident fals că fiecare experiență poate fi individualizată în mod unic prin caracterul său calitativ distinctiv sau *quale.* Nu trebuie să confundăm natura calitativă a experienței cu natura calitativă a obiectului experienței. Aceasta din urmă, și nu prima, este cea care individualizează experiența.

4) Chiar dacă lărgim conceptul de experiență pentru a include gândirea că ceva este așa cum este sau gândirea la ceva, ceea ce diferențiază în esență gândirea unui lucru de alta nu este ceea ce simți sau cum simți atunci când gândești ceea ce gândești. A gândi că 2 + 2 = 4 este diferit de a gândi că 25 x 25 = 625, iar ambele diferă de a gândi că democrații vor câștiga următoarele alegeri.17 Ele diferă în măsura în care obiectele lor sunt cele care [le] specifică sau le individualizează în esență. Putem gândi *că* ceva este așa sau așa sau ne putem gândi la *un* lucru sau altul, fără a avea simultan vreun fel de atitudine afectivă ; prin urmare, nu este nevoie să existe „cum simțim” atunci când ne gândim la asta. Gândirea la lei, la Richard Inimă de Leu sau la străzile din León poate fi însoțită de o respirație leonină. Dar, contrar sugestiei lui Chalmers , specificarea aurei asociate nu înseamnă caracterizarea *a ceea ce se simte* când te gândești la acele lucruri, cu atât mai puțin caracterizarea gândului în sine pentru a-l individualiza. Faptul că cineva asociază gândirea la unul dintre aceste lucruri cu o aură leonină nu răspunde la întrebarea (ciudată) „Cum e să te gândești la lei (Richard Inimă de Leu, străzile din León)?” și cu siguranță nu distinge între a te gândi la lei și a te gândi la León sau la Richard I.

FUNDAMENTELE FILOZOFICE ALE NEUROȘTIINȚEI

Extras din Capitolul 14: Observații finale

MAXWELL BENNETT și PETER HACKER

14.5 DE CE ESTE IMPORTANT

În ceea ce privește întrebarea cum va afecta următorul experiment

Imaginați-vă un om de știință care citește discuțiile noastre analitice cu un anumit grad de perplexitate. Poate fi moderat interesat de unele dintre analizele noastre generale, dar în același timp dezorientat de ceea ce pare a fi un flux nesfârșit de deconstrucție logică. După ce citesc expunerile noastre inițiale, s-ar putea întreba: „Chiar sunt importante toate acestea?". „La urma urmei", ar putea spune, „cum va afecta acest lucru următorul experiment ?". Avem încredere că orice cititor care ne-a urmărit până aici nu va fi tentat să pună această întrebare, deoarece aceasta dezvăluie o lipsă de înțelegere.

Nu este treaba noastră dacă reflecțiile analitice vor afecta sau nu experimentele viitoare . Acestea pot afecta sau nu - depinde de experimentul în cauză și de presupunerile neurocercetătorului. Din cele spuse până acum, ar trebui să fie evident că, dacă argumentele noastre sunt convingătoare, cea mai bună cale de acțiune în cazul unor experimente este abandonarea lor. [1] Alții

Ar trebui să le reproiecteze. [2] Este posibil ca majoritatea să nu fie afectate, deși ipotezele inițiale vor trebui probabil reformulate, iar rezultatele descrise foarte diferit. [3]

Ceea ce ne interesează este să înțelegem cel mai recent experiment

Interesul nostru nu constă în proiectarea următorului experiment, ci în *înțelegerea* ultimului. Mai general, investigațiile conceptuale contribuie în primul rând la înțelegerea a ceea ce se știe și la claritatea formulării întrebărilor despre ceea ce este necunoscut. Nu ar avea nicio importanță dacă reflecțiile noastre *nu ar avea niciun* efect asupra următorului experiment, dar ele au un efect considerabil asupra interpretării rezultatelor experimentelor anterioare. Și nu există nicio îndoială că ele au ceva de contribuit la formularea întrebărilor, la formularea acestora și la distingerea între întrebările semnificative și cele confuze. (Dacă avem dreptate, atunci întrebările despre „problema integrării ", înțeleasă ca problema modului în care creierul formează imagini, sunt în mare măsură expresii ale confuziei, [iar] o mare parte din dezbaterea despre imaginile mentale se bazează pe o neînțelegere. [)]

Contează? Da, dacă înțelegerea contează.

importantă toată această aparentă disecție logică, toată această examinare detaliată a cuvintelor și a utilizării lor ? Chiar *are nevoie neuroștiința de toate acestea? Dacă forța motrice din spatele eforturilor neuroștiințifice este dorința* de a înțelege fenomenele neuronale și relația lor cu capacitățile psihologice și exercitarea lor, atunci contează foarte mult. Pentru că, indiferent de strălucirea experimentelor neurocercetătorilor și de rafinamentul tehnicilor lor , dacă există confuzie conceptuală în întrebările lor sau erori conceptuale în descrierea rezultatelor cercetării lor, atunci aceștia nu au înțeles ceea ce și-au propus să înțeleagă.

Mulți neurocercetători contemporani care lucrează în domeniul neuroștiinței cognitive

sunt de acord că apărarea de către Sir John Eccles a unei forme de dualism a fost o greșeală - iar la baza erorii lui Eccles se află o *confuzie conceptuală* . Referindu-ne la diverse teorii ale unor distinși neurocercetători cognitivi contemporani, am încercat să arătăm că, departe de a fi eradicată prin respingerea superficială a diferitelor forme de dualism cartezian, eroarea conceptuală rămâne larg răspândită. Ea afectează și infectează plauzibilitatea întrebărilor puse, caracterul experimentelor concepute pentru a răspunde la acestea, inteligibilitatea descrierilor rezultatelor acestor experimente și coerența concluziilor trase din ele. Și acest lucru este, fără îndoială, important, atât pentru înțelegerea a ceea ce *au realizat* neurocercetătorii actuali, cât și pentru progresul viitor al neuroștiinței cognitive.

De ce este important pentru publicul educat?

Acest lucru este, de asemenea, foarte important pentru publicul educat, deoarece, indiferent dacă anumiți neurocercetători se înșală sau nu, nu există nicio îndoială că metodele descriptive pe care le folosesc derutează publicul nespecializat. Este de înțeles că neurocercetătorii sunt dornici să comunice cunoștințele despre funcția creierului pe care le-au dobândit în ultimele decenii și să împărtășească cu publicul educat pasiunea pe care o simt pentru subiectul lor de studiu. Acest lucru este demonstrat de avalanșa de cărți scrise de numeroși membri distinși ai profesiei. Dar atunci când vorbesc despre gândirea și raționamentul creierului, despre una dintre emisferele sale care știe ceva ce nu informează cealaltă, despre creierul care ia decizii fără ca persoana să fie conștientă de acest lucru, despre imagini mentale care se rotesc în spațiul mental etc., neurocercetătorii promovează o formă de mistificare, cultivând o neuromitologie deplorabilă în toate sensurile. Pentru că, în primul rând, procedând astfel, fac orice altceva decât să genereze înțelegere în rândul publicului nespecializat căruia i se adresează. În al doilea rând, publicul larg va căuta în neuroștiințe răspunsuri la pseudo-întrebări pe care nu ar trebui să le pună și la care neuroștiința nu poate răspunde. Și de atunci încolo, acest public dezamăgit va ignora întrebările cu adevărat importante pe care neuroștiința le *poate* pune și le poate răspunde . Iar acest lucru este, fără îndoială, important.

Despre nevoia de claritate conceptuală

De-a lungul acestei cărți, am încercat să demonstrăm că claritatea structurilor conceptuale este la fel de importantă pentru neuroștiința cognitivă ca și claritatea metodelor experimentale . Claritatea conceptuală scoate la iveală, mai degrabă decât împiedică, marile contribuții ale neuroștiinței cognitive la înțelegerea noastră asupra rădăcinilor biologice ale abilităților umane. Într-adevăr, realizările neuroștiinței pot fi pe deplin apreciate doar atunci când umbrele aruncate de confuzia conceptuală sunt risipite.

NEUROȘTIINȚĂ ȘI FILOZOFIE

MAXWELL BENNETT

O ODISEE PERSONALĂ

Când un potențial de acțiune propagator ajunge la o sinapsă la terminalul axonal al unui neuron presinaptic, simbolizat de triunghiul mic din Figura 1, acesta induce eliberarea de molecule neurotransmițătoare, așa cum se arată în imaginea inserată a unei sinapse din partea stângă jos a Figurii 1. Transmițătorul difuzează printr-o fantă îngustă și se leagă de receptorii de pe membrana postsinaptică. Această legare induce deschiderea de canale și, adesea, generarea de potențiale de acțiune în neuronul postsinaptic. Pentru acest proces sunt necesare câteva sute de proteine (Sieburth și colab., 2005). Am petrecut peste 40 de ani investigând mecanismele implicate în transmiterea la sinapsă (Bennett, 2001) și am inițiat recent o serie de studii asupra modului în care rețelele sinaptice își îndeplinesc funcțiile în creier. Aceste rețele, formate din miliarde de neuroni, fiecare cu până la zece mii de sinapse, pot fi găsite în părțile creierului care trebuie să funcționeze normal pentru a-și aminti un eveniment nou mai mult de un minut (hipocampul), pentru a vedea (retina și corneea).

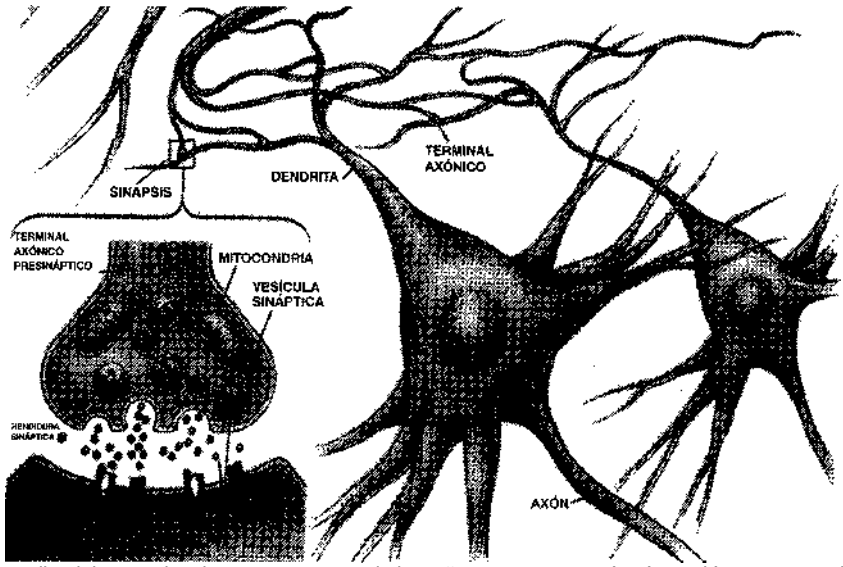


FIGURA 1: În axonii a doi neuroni au loc procese care culminează cu un contact strâns între ei în puncte numite sinapse. O sinapsă este prezentată în imaginea inserată, mărită în colțul din stânga. Textul explică funcția acestei sinapse .

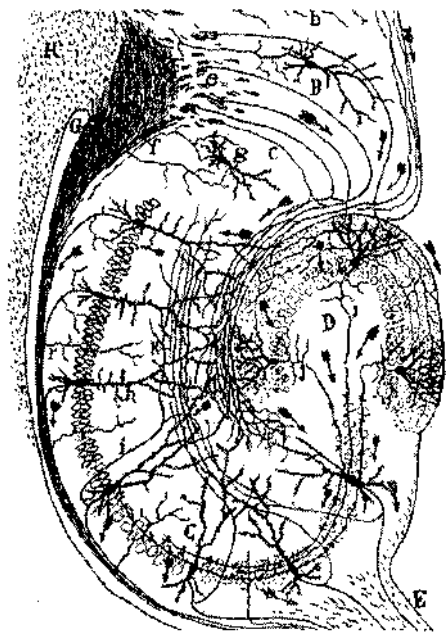
cortexul vizual primar) și pentru a dobândi o varietate de abilități motorii (cerebel).

Rețelele sinaptice pe care le-am ales pentru studiul inițial au fost cele ale hipocampului (Bennett, Gibson și Robinson, 1994). Distribuția generală a tipurilor de neuroni hipocampici și a sinapselor acestora, descrisă pentru prima dată de Ramón y Cajal (1904), este prezentată în Figura 2. Literele din figură se referă la diferite părți ale hipocampului, precum și la tipurile de neuroni (indicate prin elipsoide negre, cu axoni și dendrite lungi și înguste), împreună cu conexiunile lor sinaptice, așa cum se menționează în legenda figurii. O abordare inginerească pentru descrierea funcției hipocampice - necesită dezvoltarea unei reprezentări a rețelei neuronale precum cea prezentată în dreapta Figurii 2 și descrisă în

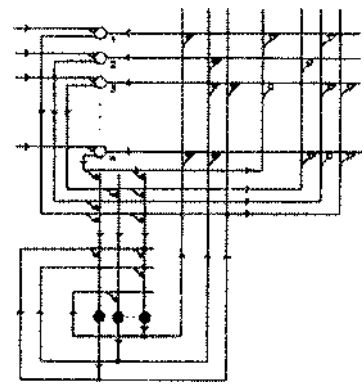
legendă. Tipurile de neuroni sunt indicate aici prin cercuri ; dendritele și axonii lor prin linii drepte; iar sinapsele prin dreptunghiuri mici. În tinerețe, după ce tocmai mi-am obținut diploma în inginerie electrică, am fost intrigat de sugestia lui Brindley (1967) că unele dintre sinapsele din rețele precum acestea erau modificabile. Ceea ce a vrut el să spună prin acest termen este că sinapsele sunt permanent susceptibile la schimbarea proprietăților lor ca o consecință a sosirii unui potențial de acțiune la terminalul axonului. Având în vedere această posibilitate, ar putea fi ca „mecanismele de memorie și condiționare ale sistemului nervos să stocheze informații prin intermediul sinapselor modificabile". Mai târziu, Brindley însuși (1969) a demonstrat în modele de rețele neuronale cum astfel de sinapse modificabile ar putea „îndeplini multe sarcini simple de învățare ". Ulterior, Marr (1971), care a lucrat și la Universitatea din Cambridge, a remarcat că „cea mai importantă caracteristică a arhicortexului (hipocampului) este capacitatea sa de a îndeplini un tip simplu de sarcină de memorare". El a fost primul care a sugerat că o rețea colaterală recurentă (Fig. 2) ar putea acționa ca o memorie asociativă dacă eficacitatea sinapselor excitatorii ar fi modificabilă și dacă potențialele de membrană ale neuronilor piramidali ar fi stabilite de interneuroni inhibitori care calculează activitatea generală a rețelei. Sugestia sa a fost formulată în termeni inginerești pe care eu și mulți alții i-am găsit foarte atractivi pentru studii teoretice și experimentale ulterioare . Colegii mei și cu mine am urmat această abordare conceptuală generală a lui Brindley și Marr și am identificat condițiile în care o reprezentare a hipocampului în rețea neuronală, cum ar fi cea din Figura 2, ar putea funcționa (Bennett, Gibson și Robinson, 1994). Sugerăm că „recuperarea unei amintiri începe cu activarea unui set de neuroni piramidali care se suprapune cu memoria care trebuie

recuperată" și că „activarea diferitelor seturi de neuroni piramidali are loc apoi în etape discrete, sincrone" până când modelul neural al memoriei stocate este recuperat (Fig. 2).[1]

Cu toate acestea , există două aspecte ale acestui tip de abordare inginerească pentru descrierea funcționării rețelelor sinaptice și, în consecință , a creierului, pe care le consider înșelătoare și care sunt detaliate în secțiunile următoare.[2]



—Sinapse excitatorii ineficiente Sinapse excitatorii eficiente Sinapse inhibitorii



Q Neurona excitatoria

Neurona inhibidora

FIGURA 2: Desenul din stânga reprezintă neuronii hipotalamusului și conexiunile lor sinaptice, așa cum sunt desenate în cerneală de Ramón y Cajal (Figura 479, în Ramón y Cajal, 1904). Literele din figură se referă la diferitele componente ale hipocampului relevante pentru scopurile noastre: D, girusul dentat; K, colateralele recurente ale celulelor piramidale din regiunea CA3 a hipocampului (C), care le permit acestora să formeze sinapse între ele și, de asemenea, să se proiecteze pentru a forma sinapse cu neuronii piramidali din regiunea CAT a hipocampului (h).

În dreapta este o diagramă a rețelei sinoptice de bază a regiunii CA3 a hipocampului , compusă din neuroni piramidali (cercuri goale) și interneuroni inhibitori (cercuri umplute). Neuronii piramidali stabilesc conexiuni aleatorii între ei prin intermediul colateralelor lor recurente. Înainte de învățare, aceste conexiuni sunt ineficiente. [PRIMA SURSĂ MAJORĂ DE INTERES PRIVIND ÎNȚELEGEREA FUNCȚIEI REȚELELOR CELULARE DIN CORTEXUL CEREBRAL]

(Continúa)

Având în vedere că știm că o leziune la nivelul hipocampului împiedică o persoană să își amintească un eveniment mai mult de aproximativ un minut, cum identificăm celulele din hipocampul normal care sunt implicate în capacitatea noastră de a ne aminti, care sunt conexiunile sinaptice dintre aceste celule și cum funcționează aceste sinapse? De exemplu, în orice volum dat al hipocampului, există multe cazuri; după învățare, un subset al acestora devine eficient, iar în starea finală de dezvoltare a rețelei, există conexiuni sinaptice excitatorii ale căror puncte forte sunt considerate a fi egale cu unu (triunghiuri goale), iar altele ale căror puncte forte au devenit zero (cercuri goale). Interneuronii inhibitori primesc conexiuni aleatorii de la multe celule piramidale și, de asemenea, de la neuroni inhibitori . Neuronii inhibitori, la rândul lor, proiectează către neuronii piramidali. Puterea oricărei sinapse care implică interneuroni inhibitori este considerată a fi fixă. Starea inițială a sistemului este determinată de un model de activare care ajunge la neuronii piramidali din axonii fibrelor mușchioase care ies din regiunea D din diagrama lui Ramón y Cajal sau din calea perforantă directă situată deasupra regiunii D - un model reprezentat de liniile care intră din stânga. Odată ce starea inițială este stabilită, sursa externă este eliminată, iar rețeaua recurentă CA3 își actualizează starea internă ciclic și sincron.

În modelul rețelei neuronale, neuronii inhibitori sunt reprezentați ca dispozitive liniare cu acțiune rapidă, care produc ieșiri proporționale cu intrările lor și joacă un rol reglator important în stabilirea potențialelor de membrană ale neuronilor piramidali. Probabilitatea ca un neuron să se excite într-o așa-numită memorie stocată, care determină numărul mediu de neuroni activi atunci când o memorie este recuperată, poate fi setată după bunul plac. Se presupune că „amintirile" acestei rețele sunt stocate în sinapse colaterale recurente folosind o diagramă Hebbian cu două valori. Teoria ia în considerare atât corelațiile spațiale dintre punctele forte învățate ale sinapselor colaterale recurente, cât și corelațiile temporale dintre starea rețelei și aceste puncte forte sinaptice. Se presupune că recuperarea memoriei începe cu activarea unui set de neuroni piramidali CA3 care se suprapun cu memoria care urmează să fie recuperată, precum și cu activarea unui set de neuroni piramidali care nu sunt implicați în memoria care urmează să fie recuperată ; activarea ambelor seturi de neuroni este probabil indusă de sinapsele pe care axonii căii perforante le formează pe neuronii $_{CA3}$. Activarea diferitelor seturi de neuroni piramidali se desfășoară apoi în etape sincrone discrete (pentru detalii, vezi Bennett, Gibson și Robinson, 1994).

Există mai multe celule gliale decât neuroni, iar aceste celule gliale sunt de diverse tipuri, așa cum este cazul neuronilor. Descoperirea undelor care se propagă și transmit activitatea între celulele gliale astrocitare (Cornell-Bell și colab., 1990), deși mult mai lent decât între neuroni (Bennett, Farnell și Gibson, 2005), introduce o complexitate considerabilă în căutarea corelațiilor celulare ale memoriei în hipocampus. Deși aceste unde gliale au fost considerate irelevante pentru căutarea corelațiilor celulare ale atributelor noastre psihologice (Koch, 2004), nu au fost efectuate experimente pentru a demonstra în mod concludent acest lucru. Având în vedere proximitatea fizică dintre celulele gliale și neuroni, va fi dificil de determinat dacă primele sunt relevante în căutarea corelațiilor celulare, chiar dacă putem manipula genetic proteinele acestor tipuri diferite de celule. Cred că ar fi mult mai sensibil să căutăm corelații „celulare", decât pur și simplu „neuronale". În plus, abia începem să înțelegem vasta gamă de relații sinaptice dintre neuroni, dintre celulele gliale și dintre ambele tipuri de celule, precum și diversitatea mecanismelor sinaptice specializate care operează între ele. În opinia mea, divinizarea actuală a neuroștiinței este deplasată , după cum reiese, de exemplu, din presupunerea că rețele precum cele din Figura 2 oferă indicii importante despre funcționarea rețelelor sinaptice în orice parte a creierului. Pentru a oferi dovezi în acest sens, voi prezenta acum câteva exemple despre cât de dureros de lent este progresul în înțelegerea chiar și a rețelelor sinaptice relativ simple care operează în retină, cortexul vizual primar și cerebel.

Rețelele retiniene

Cea mai simplă și mai accesibilă parte a sistemului nervos central este retina , care în timpul dezvoltării apare inițial ca o evaginație a creierului. Din acest motiv, cel mai important histologist al sistemului nervos, Ramón y Cajal, a considerat retina locul ideal pentru a începe cercetările privind înțelegerea funcționării rețelelor sinaptice ale sistemului nervos central. La scurt timp după ce am început această cercetare, Barlow și Levick (1965) au descoperit ceea ce sunt cunoscuți sub numele de neuroni ganglionari direcțional selectivi în retina unor specii. Un neuron ganglionar direcțional selectiv este acela care trimite impulsuri de mare viteză atunci când un obiect se mișcă într-o anumită direcție (de unde și numele de direcție preferată) în fața fotoreceptorilor bastonaș sensibili la lumină conectați la celula ganglionară, situată deasupra acesteia. Când obiectul se mișcă în direcția opusă, neuronul ganglionar nu este activat (și, în consecință, acea direcție se numește direcție nulă). Barlow și Levick au propus schema prezentată în Figura 3 (jos) pentru a explica

originile reticulare ale selectivității direcționale. În această rețea, un semnal inhibitor compensator blochează semnalul excitator pentru mișcarea în direcția nulă. Un interneuron lateral transmite un semnal inhibitor în direcția nulă, dar nu în direcția preferată, în timp ce semnalul excitator acționează local (Fig. 3). Prin urmare, în mișcarea în direcția nulă, inhibiția ajunge înainte de excitație și poate interacționa cu aceasta, dar în mișcarea în direcția preferată, inhibiția ajunge mai târziu decât excitația. Acest algoritm este explicat în textul care însoțește Figura 3. Această rețea simplă m-a impresionat foarte mult când a fost publicată în 1967. A fost prima analiză de acest fel care a oferit o explicație a funcționării unei rețele sinaptice reale.

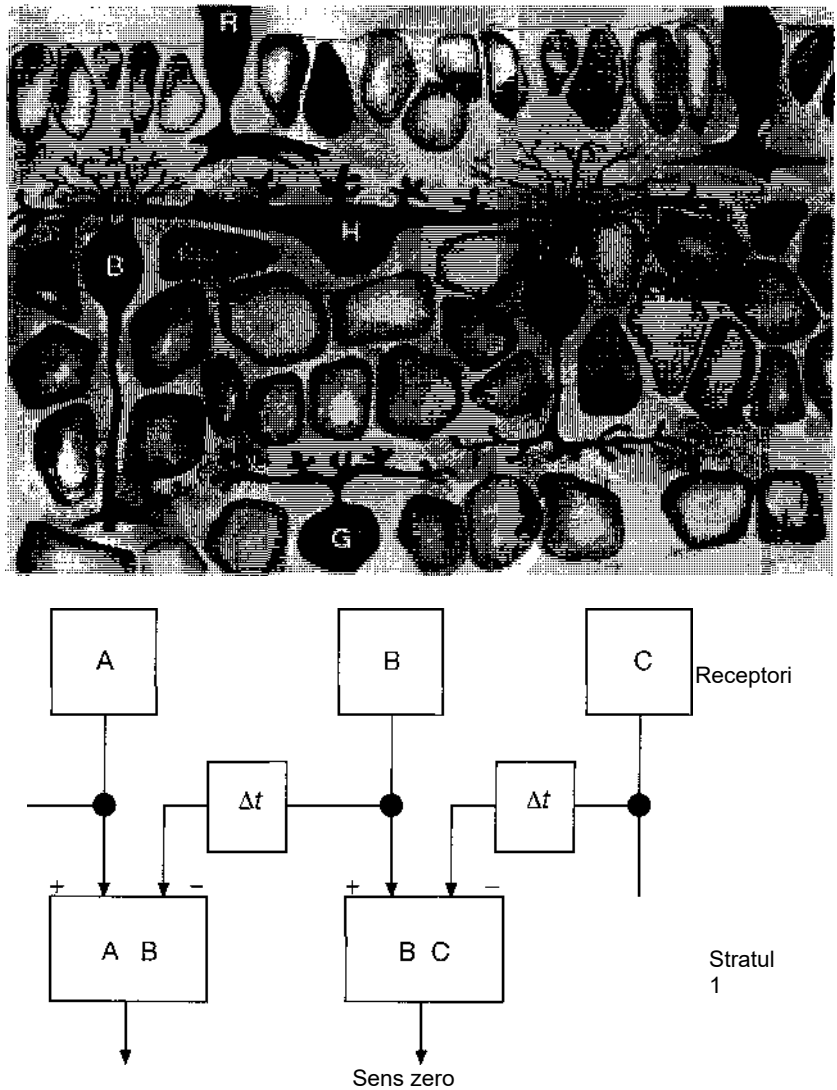


FIGURA 3: Graficul de sus este o diagramă a unora dintre principalele celule ale retinei identificate în 1965. R, fotoreceptor cu tijă; B, neuron bipolar; H, neuron orizontal ; G, neuron ganglionar care conectează retina la creier.

Diagrama de mai jos prezintă algoritmul propus de Barlow și Levick (1965) pentru a explica selectivitatea direcțională a neuronilor ganglionari . A, B și C sunt receptori care pot răspunde la un obiect care se mișcă deasupra lor în direcția nulă.

(Continúa)

Dacă acest algoritm este corect, atunci se pune întrebarea despre identitatea interneuronului lateral *şi* a conexiunilor sale sinaptice. Cât de mult au avansat neurologii în ultimii patruzeci de ani, de când a fost sugerat algoritmul, în identificarea componentelor celulare care efectuează calculele necesare și a conexiunilor acestora? În momentul cercetării inițiale privind selectivitatea direcțională, erau cunoscute aproximativ zece tipuri diferite de celule din retina vertebratelor (Ramón y Cajal, 1904; Polyak, 1941). Astăzi, sunt recunoscute cel puțin cincizeci de tipuri diferite de celule (Masland, 2001), fără a lua în considerare diferitele tipuri de celule gliale care intră în contact strâns cu neuronii, dar nu transmit potențiale de acțiune. Au fost necesari patruzeci de ani de cercetare pentru a identifica unele dintre mecanismele celulare retiniene responsabile de selectivitatea direcțională, iar multe întrebări importante rămân fără răspuns (Fried, Munch și Werblin, 2005). Unele dintre caracteristicile specificate în schemă de Barlow și Levick (1965) au fost identificate într-un neuron numit celulă amacrină stelată (Fried, Munch și Werblin, 2002). Cu toate acestea, schema a trebuit să fie modificată radical după descoperirea faptului că inputurile celulelor ganglionare sunt ele însele selective direcțional (Vaney și Taylor, 2002).

(indicat printr-o săgeată), în direcția preferată. Fiecare dintre acești receptori poate excita activitatea unităților imediat de sub el. Fiecare casetă care conține un At este o unitate care, dacă este excitată de receptorul conectat la aceasta, va împiedica, după o întârziere de lungime At, excitarea unității adiacente în direcția nulă. Când un obiect se mișcă în direcția nulă, activitatea electrică a unui receptor excitat (de exemplu, C) excită (+) o unitate din stratul imediat de sub el și, în același timp, inhibă (-) următoarea unitate în direcția nulă; la rândul său, fiecare receptor, în special C, B și A, efectuează acest proces pe măsură ce obiectul se mișcă deasupra lor. Unitățile de întârziere (etichetate At) asigură că procesul inhibitor oprește activitatea excitatorie a lui A și B care circulă prin aceste porți dacă mișcarea este în direcția nulă, dar ajunge la porți prea târziu pentru a produce o astfel de inhibare dacă mișcarea este în direcția preferată.

Prin urmare, se pare că circuitele retiniene au mai multe niveluri care determină selectivitatea direcțională a celulelor ganglionare , implicând probabil activitatea a cel puțin patru rețele sinaptice retiniene diferite și specifice, care nu au fost încă identificate (Fried, Munch și Werblin, 2005). Acest lucru ilustrează dificultățile în înțelegerea unei proprietăți de rețea chiar și a ceea ce a fost considerată cea mai

simplă și mai accesibilă parte a sistemului nervos central.

Rețelele cortexului vizual primar

Cea mai interesantă conferință științifică la care am participat vreodată a avut loc acum mai bine de treizeci de ani, la Simpozionul de la Cold Spring Harbor privind sinapsele, în 1975. James Watson m-a invitat să prezint descoperirile noastre privind plasticitatea sinapselor dintre nervi și celulele musculare, în special descoperirile noastre privind dezvoltarea punctului inițial de contact dintre terminalul nervos și mușchi, modul în care acest punct devine o enclavă de terminale nervoase în exces și modul în care toate aceste terminale, cu excepția uneia, sunt eliminate în timpul procesului (Bennett și Pettigrew, 1976). Înainte de a ajunge la simpozion, nu știam că, în urma prezentării mele, Hubei, Wiesel și LeVay (1976) aveau să ofere o explicație excelentă a dezvoltării conexiunilor de rețea sinaptică în cortexul vizual primar (Vi), care stau la baza formării coloanelor de neuroni dominate de conexiuni către un ochi sau altul. În stadiile incipiente ale dezvoltării, acești neuroni au conexiuni către ambii ochi, dar după un proces de eliminare sinaptică, unul sau altul devine în cele din urmă dominant în ceea ce privește conectivitatea. Hubei, Wiesel și LeVay au demonstrat plasticitatea considerabilă a acestui proces într-un mod cu adevărat spectaculos, arătând cum, dacă vederea este limitată la un singur ochi în timpul dezvoltării timpurii, celălalt ochi domină conectivitatea sinaptică. Într-o perioadă critică de dezvoltare, este posibil să se inverseze acest rezultat dacă vederea este restabilită la ambii ochi.

La cincisprezece ani după acea prezentare memorabilă, la Simpozionul de la Cold Spring Harbor din 1990, Wiesel și colegii săi, Gilbert și Hirsch, au reluat subiectul plasticității conexiunilor sinaptice în cortexul vizual primar (V). Cu toate acestea, de data aceasta, au mutat accentul pe gradul de plasticitate al rețelei sinaptice din cortexul vizual adult. Cercetările lor, bazate pe tehnici de înregistrare electrofiziologică, au sugerat că, în câteva luni de la o leziune retiniană, are loc o

reorganizare la scară largă a rețelelor sinaptice din cortex (Gilbert, Hirsch și Wiesel, 1990). Această lucrare a fost interpretată ca însemnând că rețelele sinaptice corticale adulte se pot reajusta după pierderea unui semnal senzorial de intrare (Gilbert, 1998). O înțelegere clară a mecanismelor de plasticitate responsabile de aceste reajustări este de o importanță considerabilă dacă vrem să-i ajutăm pe cei care au nevoie de un tratament de reabilitare adecvat.

Între două și șase luni după leziunile retiniene binoculare care privează o zonă din cortexul vizual primar (VJ) de inputul său normal , se detectează o activitate recurentă direcționată către stimuli până la 5 mm în zona cortexului lipsită de inputul său vizual (Gilbert, 1998). Se crede că modificări mai modeste ale topografiei corticale, cuprinse între 1 și 2 mm, apar imediat (în câteva minute până la ore) după astfel de leziuni (Gilbert și Wiesel, 1992). La cincisprezece ani după aceste studii privind plasticitatea cortexului vizual adult, Logothetis și colegii săi au studiat semnalele din cortexul vizual primar al macacului după...

Au fost create mici leziuni retiniene binoculare pentru a determina amploarea *și* evoluția în timp a reorganizării cortexului vizual (Vp Smirnakis et al., 2005). Leziunile retiniene au fost induse prin fotocoagulare cu laser, creând un scotom, sau o breșă perceptivă, cu diametrul de 4-8 grade în câmpul vizual . Aceste leziuni au privat o parte a cortexului de input vizual de la fiecare ochi, un proces despre care se crede că maximizează reorganizarea. Regiunea cortexului vizual lipsită de input retinian se numește zona de proiecție a leziunii. Imagistica prin rezonanță magnetică funcțională (fMRI) a fost utilizată pentru a detecta modificări ale topografiei corticale a cortexului vizual la macac după aceste leziuni retiniene binoculare. Spre deosebire de studiile menționate anterior care au utilizat electrofiziologia, câmpul vizual larg oferit de imaginile fMRI a demonstrat că, în timpul celor 7,5 luni care au urmat leziunilor retiniene, cortexul vizual primar nu se apropie de capacitatea sa de răspuns normală, iar topografia sa rămâne neschimbată (Fig. 4). Profitând de câmpul vizual larg oferit de aceste imagini, electrozii de înregistrare electrofiziologică au putut fi plasați precis în zona de proiecție a leziunii. Acești electrozi au confirmat rezultatele RMN-ului funcțional. Astfel, conform observațiilor lui Logothetis și colegilor săi, utilizând două tehnici diferite, cortexul vizual primar are un potențial limitat pentru reorganizarea rețelei sinaptice , cel puțin până la câteva luni după leziunea retiniană.

Detaliile despre ce ar fi putut merge prost în ultimii cincisprezece ani de cercetare asupra plasticității corticale sunt încă necunoscute. Este suficient să spunem că complexitatea determinării proprietăților acestei prime enclave corticale care primește informații de la retină este de așa natură încât necesită mare atenție, abilități tehnice, perspectivă teoretică și determinare. Wiesel, pre-

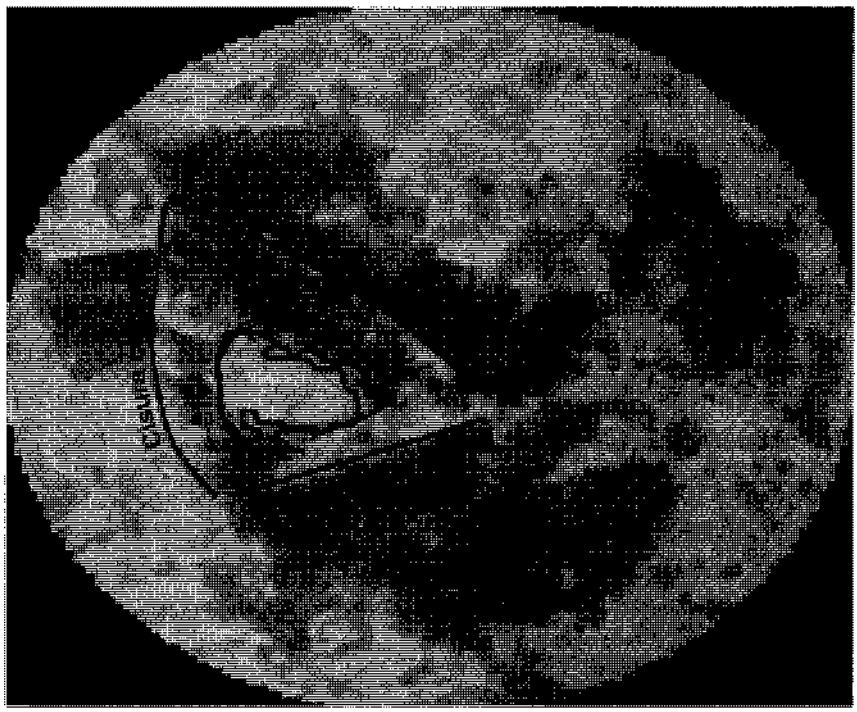


FIGURA 4: Imaginea cu contrast dependent de nivelul oxigenului din sânge (BOLD), care reflectă probabil activitatea sinoptică, măsurată prin imagistică prin rezonanță magnetică funcțională (fMRI) în interiorul unei zone de proiecție a leziunii, nu se modifică în timp după o leziune retiniană. Este prezentată o zonă cu o rază de 3 cm a cortexului vizual Vi, centrată în apropierea reprezentării foveale. Zona delimitată de șanțul calcarin, șanțul semilunar și osul occipital inferior este indicată în această secțiune aplatizată. Zonele din afara acestei zone corespund în mare parte cortexului non-vizual. Limitele zonei de proiecție a leziunii sunt prezentate la 0 zile (contur interior), 4 luni (contur exterior) și 7,5 luni (contur intermediar) după leziune. Zona de proiecție a leziunii măsurată a fost de, respectiv , 158, 179 și 180 mm² (din figura 2 din Smirnakis și colab., 2005).

Laureatul meu cu Premiul Nobel și Gilbert sunt neurocercetători de primă clasă. Dar, după cincisprezece ani de studii intensive, nu s-a ajuns la niciun consens cu privire la întrebarea foarte importantă dacă rețelele neuronale corticale vizuale primare sunt plastice (Giannikopoulos și Eysel, 2006). Am prezentat această istorie în detaliu nu pentru a da vina pe alții, ci pentru a sublinia faptul că complexitatea biologică a sistemelor reticulare sinaptice pe care încercăm să le înțelegem este considerabilă și testează abilitățile chiar și ale celor mai buni neurocercetători . Cu toate acestea, dacă nu înțelegem unele dintre proprietățile

fundamentale ale acestui prim centru de transmisie de la cortex la alte zone corticale legate de funcția vizuală, cum ar fi cele din lobul temporal a cărui funcție normală este necesară pentru recunoașterea facială, este foarte puțin probabil să înțelegem vreodată rețelele sinaptice subordonate capacității noastre vizuale în zone precum lobul temporal.

Rețelele cerebeloase

Cortexul cerebelos posedă un set de tipuri de neuroni și conexiuni sinaptice a căror aranjament și funcție par a fi deosebit de simple, așa cum se arată în Figura 5. Prin urmare, cerebelul pare a fi deosebit de potrivit pentru analiza experimentală și dezvoltarea de teorii verificabile empiric, un subiect pe care îl consider deosebit de fascinant (vezi, de exemplu, Gibson, Robinson și Bennett, 1991). Briliantul David Marr (1969) a introdus ideea că achiziția de noi abilități motorii depinde de plasticitatea sinapselor dintre terminalele axonilor paraleli și celulele Purkinje ale cerebelului (Fig. 5). Această observație s-a dovedit extrem de fructuoasă, deoarece în următorii treizeci și cinci de ani plasticitatea acestor sinapse a fost utilizată pentru a explica achiziția și reținerea pe tot parcursul vieții a multor abilități motorii, inclusiv sincronizarea motorie învățată și adaptarea reflexă (Ito, 2001). Un număr mare de studii au fost compilate asupra modelului de rețea sinaptică în care activitatea terminalelor axonilor fibrelor cățărătoare de pe celulele Purkinje se re-

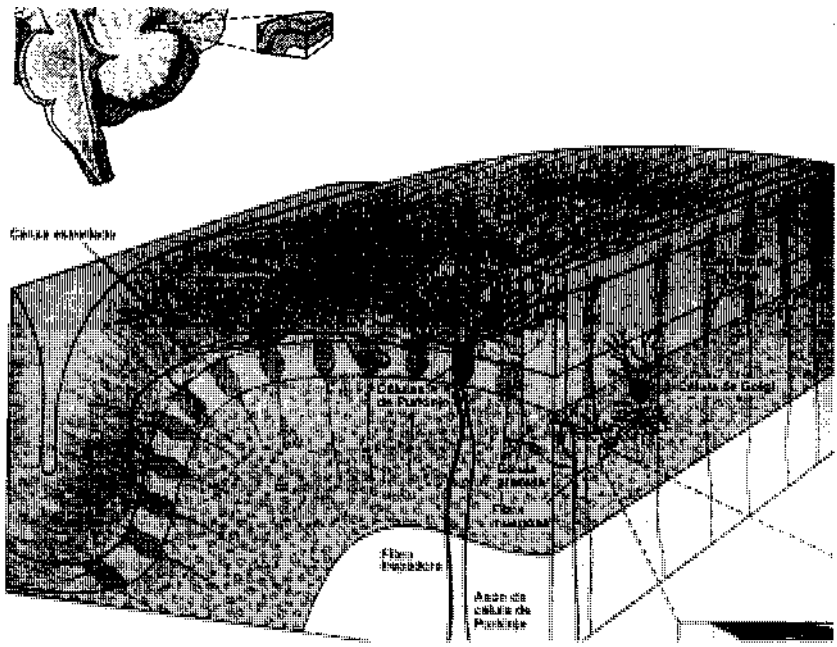


FIGURA 5: Cortexul cerebelos. Sunt prezentate celulele Purkinje mari, cu arborii lor dendritici enormi. Fiecare dintre acestea primește conexiuni sinaptice de la terminalele axonale ale fibrelor paralele derivate din celule granulare foarte mici, care, la rândul lor, primesc sinapse de la terminalele axonale ale fibrelor mușchioase. În mod special, celulele Purkinje primesc, de asemenea, conexiuni sinaptice de la un singur axon al unei fibre cățărătoare. Acest circuit este aranjat într-un model frumos, repetitiv și regulat, ceea ce face din cortexul cerebelos un subiect ideal pentru cercetarea experimentală. Sunt prezentați și neuronii inhibitori (în special, celulele coș, celulele stelate și celulele Golgi).

Intensitatea potențialelor sinaptice la terminalele axonale ale fibrelor paralele din aceleași celule este redusă atunci când ambele intrări (adică cele de la fibrele cățărătoare și cele de la fibrele paralele) sunt stimulate simultan (Fig. 5). Se spune că această formă de plasticitate sinaptică stă la baza dobândirii de noi abilități motorii. Dacă stimularea electrică este utilizată pentru a stimula direct și simultan terminalele axonale ale fibrelor cățărătoare și pe cele ale fibrelor paralele, amplitudinea potențialelor sinaptice activate de terminalele axonale ale fibrelor paralele din celulele Purkinje este redusă (Ito și Kano, 1982). Această depresie a potențialelor sinaptice necesită împerecheri repetate ale intrărilor sinaptice de la

terminalele axonale ale fibrelor cățărătoare și paralele. Depresia persistă timp de multe ore după finalizarea acestui protocol de stimulare și se numește depresie pe termen lung. În cei douăzeci și trei de ani de la descoperirea depresiei pe termen lung de către Ito și colegii săi, s-au efectuat numeroase cercetări pentru a afla baza sa moleculară (Ito, 2002).

Cu toate acestea, un experiment realizat anul trecut a provocat o mare agitație demonstrând că depresia pe termen lung nu afectează învățarea motorie la nivelul cerebelului. Llinas și colegii săi (Welsh și colab., 2005) au prevenit farmacologic depresia pe termen lung a sinapselor formate de terminalele axonilor fibrelor paralele de pe celulele Purkinje după stimularea simultană atât a terminalelor axonilor fibrelor de cățărare, cât și a celor paralele (Fig. 6). Acest lucru nu a avut niciun efect asupra dobândirii abilităților motorii implicate în testul wheelie (Lalonde, Bensoula și Filali, 1995), nici asupra dezvoltării sincronizării motorii în timpul condiționării reflexe de clipire (McCormick și Thompson, 1984). După - treizeci și șase de ani de cercetare, rețelele sinaptice și mecanismele moleculare implicate în învățarea motorie la nivelul cerebelului rămân neclare.

Complexitatea cercetării asupra sistemului nervos central

Cele trei exemple de mai sus arată progresul relativ lent al neuroștiinței în încercarea sa de a descifra chiar și complexitatea

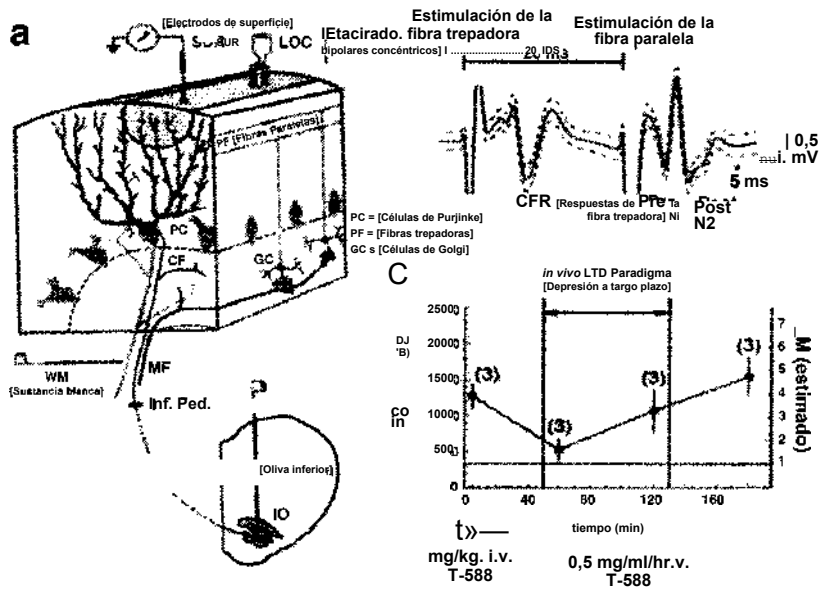


FIGURA 6: Medicamentul T-588 previne depresia pe termen lung a fibrei paralele cu sinapsa celulelor Purkinje, indusă de stimularea simultană *in vivo* a fibrei cățărătoare și a fibrei paralele. Concentrații similare de T-588 în creierul șoarecilor și șobolanilor activi nu au afectat nici învățarea motorie în testul roții de rostogolire, nici învățarea sincronizării motorii în timpul condiționării reflexe clasice de clipire . Astfel, depresia pe termen lung a fibrelor paralele cu celulele Purkinje sub controlul fibrelor cățărătoare nu este necesară pentru adaptarea motorie sau învățarea sincronizării răspunsului în două modele comune de învățare motorie care implică cerebelul. În a), sunt arătați electrozi stimulatori bipolari concentrici poziționați pe suprafața cerebeloasă (LOG) pentru a stimula un fascicul de fibre paralele (PF) și în substanța albă cerebeloasă (WM) pentru a stimula fibrele cățărătoare (CF). Un electrod de suprafață cu disc de argint (SUR) a fost utilizat pentru a înregistra potențialele de câmp evocat în celulele Purkinje (PC). Stimularea directă a olivei cerebeloase inferioare (IO) a fost utilizată pentru a verifica dacă potențialele generate de electrodul WM erau răspunsuri ale fibrelor cățărătoare (CFR). Figura b) prezintă CFR-urile și răspunsurile PF presinaptice (Nf) și postsinaptice (N2) generate de stimularea comună a FC și PF utilizând un interval de 20 m între stimuli pentru a genera un răspuns PF N2 cu depresie pe termen lung (LTD) . Liniile punctate indică o deviație standard. Figura c) prezintă rezultatele diferitelor concentrații de T-588 în creier. Sunt evidențiate patru momente diferite de injectare intravenoasă continuă de T-588 cu paradigma depresiei pe termen lung (LTD), situate între 50 și 130 m după începerea injectării. Linia orizontală de la 1 p,M indică concentrația de T-588 care a prevenit LTD *in vitro*. Numărul de creiere din probă la fiecare moment este indicat între paranteze (preluat din Figura 1 a lui Welsh et al., 2005).

a părților „mai simple" ale sistemului nervos central. Aceștia sugerează că este înțelept să fim precauți cu privire la auto-deificarea care urmează convingerii că neurocercetătorii înțeleg multe dintre funcțiile sistemului nervos central (dacă înțeleg vreuna) și subliniază cât de mult este necesar să ne oprim și să reflectăm

înainte de a accepta multe dintre afirmațiile despre ceea ce fac rețelele sinaptice ale creierului.

A DOUA SURSĂ PRINCIPALĂ DE INTERES LEGATĂ DE ÎNȚELEGEREA NOASTRĂ A ROLULUI REȚELELOR CELULARE DIN REGATUL UNIT DIN SCOATĂ DE LA REGATUL UNIT

Am menționat două aspecte ale abordării inginerești a înțelegerii rețelelor sinaptice care merită reflecție. Am luat în considerare primul aspect în a doua secțiune a acestui capitol, și anume dificultatea considerabilă în determinarea - proprietăților biologic relevante pentru construirea unui tip de rețea inginerească care poate fi utilizată pentru a înțelege mai bine rețeaua biologică. A doua dificultate apare atunci când atributele pe care le atribuim în mod normal oamenilor și, în unele cazuri, altor animale, sunt atribuite rețelelor sinaptice , fie înainte, fie după ce acestea au fost reduse la dispozitive inginerești cu grade diferite de complexitate și adaptabilitate. Se spune că anumite rețele sinaptice sau grupuri de rețele sinaptice din creier își amintesc, văd și aud. De exemplu, se observă că „putem considera toată vederea ca o căutare continuă de răspunsuri la întrebările puse de creier. Semnalele de la retină constituie «mesaje» care conțin aceste răspunsuri" (Young, 1978). Se spune că cortexul vizual al polului occipital (Fig. 7) posedă neuroni care „furnizează argumente pe baza cărora creierul își construiește ipotezele de percepție" (Blakemore, 1977). În ceea ce privește zonele creierului care trebuie să funcționeze pentru ca noi să vedem culorile, se afirmă că aceste zone sunt implicate în „interpretarea de către creier a proprietății fizice a obiectelor (reflectanța lor), interpretare care îi permite să dobândească rapid cunoștințe despre proprietatea de reflectanță" (Zeki, 1999).

Nu numai că diverse atribute psihologice sunt atribuite fasciculelor de rețele sinaptice din creier, ci sunt atribuite și emisferelor cerebrale întregi (Fig. 7). De exemplu, se sugerează că „ emisfera dreaptă este capabilă să înțeleagă limbajul,

dar nu și sintaxa” și că „capacitatea emisferei drepte de a face inferențe este extrem de limitată” (Gazzaniga, Ivry și Mangun, 2002). Se mai afirmă că „emisfera stângă poate percepe și recunoaște fețe și prezintă abilități superioare de recunoaștere - atunci când vine vorba de fețe familiare” și că „emisfera stângă adoptă o strategie cognitivă utilă în rezolvarea problemelor, dar emisferei drepte îi lipsesc abilități cognitive suplimentare” (Gazzaniga, Ivry și Mangun, 2002).

Aceste afirmații conform cărora rețelele sinaptice, fie ele biologice sau rezultatul unei reduceri utile la un dispozitiv ingineresc, posedă atribute psihologice, mi s-au părut extraordinare . Progresul lent și anevoios al neuroștiinței în utilizarea abordării inginerești pentru a ilustra rețelele sinaptice nu a implicat logic că teoria posedării de atribute psihologice de către acestea nu poate fi susținută, dar m-a determinat să nu am încredere în ea. Așa că am căutat ajutor pentru a ajunge la o clarificare conceptuală de la acei cercetători instruiți profesional în astfel de domenii, în special de la filosofi. După ce am citit lucrări ale unora dintre figurile importante—

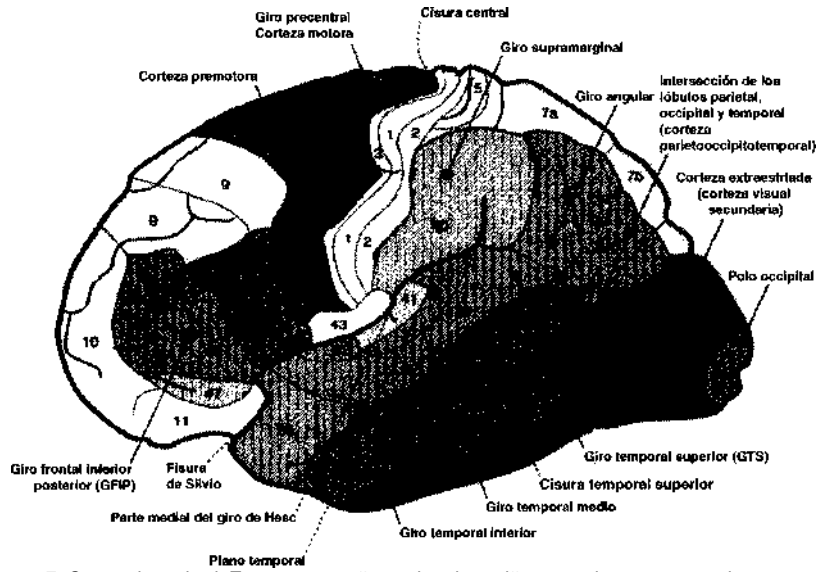


FIGURA 7: Cortexul cerebral. Este prezentată o vedere laterală cu zonele numerotate desemnate de Brodmann (1909) ca indicatori ai claselor celulare distincte și ai relațiilor dintre acestea. Rețelele celulare legate de vedere, de exemplu, sunt situate în polul occipital.

Urmând lucrările unora dintre cei mai importanți filozofi ai secolului trecut, precum Russell, Wittgenstein și Quine, am inițiat o discuție cu câțiva filozofi contemporani, în special cu Peter Hacker de la Oxford. Dialogul nostru despre posibilitatea atribuirii atributelor psihologice rețelelor sinaptice a avut loc exclusiv prin intermediul internetului și s-a încheiat înainte de a ne fi întâlnit vreodată în persoană. Pentru mine, a fost o călătorie extrem de satisfăcătoare. M-a obligat să reconsider istoria neuroștiinței de la Galen în secolul I până în prezent și să - întreprind, împreună cu Peter, o analiză critică a opiniilor figurilor de marcă din această disciplină care i-au condus pe neurocercetători la dificultățile cu care se confruntă astăzi. Rodul acestui dialog este cartea noastră, *Fundamentele filozofice ale neuroștiinței.* Satisfacția mea față de acest demers a fost sporită și mai mult de invitația din partea Asociației Americane de Filosofie de a participa la o dezbatere critică cu profesorii Dennett și Searle pe marginea tezelor cărții noastre. Rezultatul a fost o claritate sporită și, în consecință, o îmbunătățire a încercării noastre de a stabili adevărul cu privire la ceea ce neuroștiința a stabilit și pe care se poate baza

să realizeze în viitor. În acest fel, am contribuit la promovarea obiectivelor neuroștiinței de a ne ajuta să înțelegem ce înseamnă să fii om și să atenuezi suferința umană.

REFERINȚĂ BIBLIOGRAFICĂ

Barlow, H., „Doctrina neuronului în percepție", în MS Gazzaniga (comp.), *The New Cognitive Neurosciences, ediția* a 4 -a, Cambridge, MIT Press, 1997, pag. 421.

Blake, R. y NK Logothetis, «Visual competition», en *Nature Reviews Neuroscience,* nr. 3, 2002, pag. 13-21.

Crick, E. și C. Koch, „Un cadru pentru conștiință", în *Nature Reviews Neuroscience,* nr. 6, 2003, pag. 119-126.

De Lima, AD, T Voigt și JH Morrison, „Morfologia celulelor din girusul temporal inferior care se proiectează spre cortexul prefrontal la maimuța macac", în *Journal of Comparative Neurology,* nr. 296, 1990, pag. 159-172.

Koch, C., *În căutarea conștiinței,* Greenwood Village, CO, Roberts, 2004.

Leopold, DA y N. K. Logothetis, «Fenomene multistable: schimbarea viziunilor în percepție», ro *Trends in Cognitive Science,* nr. 3, 1999, pag. 254-264.

Lumer, ED, KJ Friston și G. Rees, „ Corelații neuronale ale rivalității perceptive în creierul uman", în *Science,* nr. 280, 1998, pp. 1.930-1.934.

Srinivasan, R., DP Russell, GM Edelman și G. Tononi, „Sincronizare crescută a răspunsurilor neuromagnetice în timpul percepției conștiente ", în *Journal of Neuroscience,* nr. 5.435-5.448.

Tononi, G. și GM Edelman, „Conștiință și complexitate", în *Science ,* nr. 282, 1998, pp. 1.846-1.851

REFUTĂRILE

FILOZOFIA CA ANTROPOLOGIE NAGEN

Comentariu despre Bennett și Hacker

DANIEL DENNETT

Cartea lui Bennett și Hacker , *„Philosophical Foundations of Neuroscience”* (Blackwell, 2003), o lucrare colaborativă între un filosof (Hacker) și un neurocercetător (Bennett), este o încercare ambițioasă de a redefini obiectivele cercetării în neuroștiința cognitivă, demonstrând că specialiștii în știința cognitivă și alți teoreticieni, inclusiv eu, s-au confundat reciproc folosind limbajul într-un mod sistematic „incoerent” și „confuz” din punct de vedere conceptual. Atât în stil, cât și în substanță, cartea evocă Oxfordul de la începutul anilor 1960, când filosofia limbajului obișnuit era hegemonică , iar Ryle și Wittgenstein erau autoritățile în ceea ce privește semnificațiile termenilor mentaliști sau psihologici de zi cu zi. Eu însumi sunt un produs al acelui timp și loc (așa cum este, în acest sens, Searle) și, înainte de a trece la criticile mele, care vor fi dure, vreau să subliniez ceea ce cred că este exact corect în abordarea sa: lecțiile adesea uitate ale filozofiei limbajului obișnuit.

> Studiile neuroștiințifice [...] se apropie de cele psihologice și claritatea cu privire la realizările cercetării creierului

> Aceasta presupune claritate în ceea ce privește categoriile descrierii psihologice actuale: adică categoriile de senzație și percepție, cunoaștere și memorie, gândire și imaginație, emoție și voință . În măsura în care neurologii nu disting limitele categoriilor relevante, riscă nu numai să pună întrebări greșite, ci și să interpreteze greșit propriile rezultate experimentale.
>
> (p. 115)

Acestea fiind spuse.1 Atunci când neurocercetătorii folosesc termenii comuni care constituie setul pe care l-am numit „psihologie populară", [2 trebuie] să procedeze cu maximă prudență, deoarece acești termeni poartă presupoziții de utilizare care le pot submina scopurile și pot transforma ceea ce altfel ar fi teorii promițătoare și modele empirice în nonsensuri abia deghizate. Filosoful, expert în nuanțele de sens care captivează imaginația teoreticianului, este genul potrivit de gânditor pentru a realiza acest exercițiu important de igienă conceptuală .

De asemenea, sunt de acord cu ei (deși nu aș exprima-o în termenii lor) că „dovezile pe care se bazează atribuirea atributelor psihologice altora nu sunt inductive, ci mai degrabă bazate pe judecată ; dovezile sunt logic solide" (p. 82). Acest lucru mă pune de partea lor împotriva, de exemplu, a lui [Fodor.3]

Prin urmare, sunt întru totul de acord cu presupusa motivație din spatele cărții sale. Împărtășesc, de asemenea, unele dintre principalele ținte ale criticilor sale, în special afirmația sa conform căreia neuroștiința cognitivă este plină de rămășițe carteziene care provoacă daune substanțiale. Acestea spun, de exemplu:

> În general, neurologii contemporani consideră culorile, sunetele, mirosurile și gusturile drept „constructe mentale create în creier prin procesare senzorială. Ele nu există, ca atare, în afara creierului" [citat din Kandel și colab., 1995]. Acest lucru, încă o dată, diferă de cartesianism doar prin înlocuirea minții cu creierul.
>
> (p. 113)

Aici critică un caz de ceea ce am numit „materialism cartezian "

(Consciousness Explained, 1991 [traducere în spaniolă: *La conciencia explicada,* Barcelona Paidós, 1995]) și, în opinia mea, au dreptate să creadă că mulți neurocercetători cognitivi sunt orbiți de ideea unui loc în creier (pe care îl numesc Teatrul Cartezian) unde se desfășoară un spectacol interior de construcții remarcabile în fața unei *res cogitans* (material) care constituie publicul.

Mai precis, cred că au dreptate să denunțe cartesianismul slăbit al concepției lui Benjamin Libet despre acțiunea intențională și al unora dintre lucrările teoretice ale lui Stephen Kosslyn despre imageria mentală. De asemenea, sunt de acord cu ei în a considera termenul filosofic „tehnic" *qualia deplorabil* - un dar otrăvit pentru neuroștiințe, dacă a existat vreodată unul - și împărtășesc unele dintre îndoielile lor cu privire la expresia „cum se simte", pe care Brian Farrell (1950) a fost primul care a explorat-o și pe care Thomas Nagel (1974) a făcut-o celebră. Intuiția , afirmă ei, nu este o formă de viziune interioară; nu există ochiul minții. Sunt de acord. Și a avea durere nu este același lucru cu a avea un ban; durerea *nu este ceva* care este „acolo înăuntru". Într-adevăr. Deși nu sunt de acord cu tot ce spun pe căile pe care le parcurg pentru a ajunge la toate aceste concluzii, sunt de acord cu descoperirile lor. Sau ' mai precis, sunt de acord cu concluziile mele, chiar dacă nu le citează.4

Mai mult decât incapacitatea lor de a recunoaște aceste puncte de acord cu adevărat substanțiale, sunt surprins că nucleul cărții lor, care este și nucleul atacului remarcabil de ofensiv pe care mi-l îndreaptă , este un argument pe care eu însumi l-am conceput și tratat în profunzime încă din 1969. Iată ceea ce ei numesc eroarea mereologică:

> Știm ce înseamnă pentru ființele umane să experimenteze lucruri, să vadă lucruri, să știe sau să creadă lucruri, să ia decizii, să interpreteze date ambigue, să facă conjecturi și să formuleze ipoteze. Dar știm oare ce înseamnă pentru un *creier* să vadă sau să audă, să aibă experiențe, să știe sau să creadă ceva? Avem vreo idee ce ar putea însemna pentru un *creier* să ia o decizie?

Ei răspund cu un răsunător „NU!”.

> Nu are sens să atribuim predicate psihologice (sau negările lor) creierului, decât metaforic sau metonimic. Combinația de cuvinte rezultată nu spune nimic fals; mai degrabă, nu spune absolut nimic, deoarece este lipsită de sens. Predicatele psihologice sunt predicate care se aplică în esență animalului viu în întregime, nu părților sale. Nu ochiul (și cu atât mai puțin creierul) vede, ci *noi* vedem *cu* ochii noștri (și nu cu creierul nostru, deși fără un creier cu un sistem vizual funcțional normal, nu am vedea).
>
> (p. 72)

Acest lucru este, cel puțin, foarte apropiat de argumentul pe care l-am prezentat în 1969, când am făcut distincția între nivelurile de explicație personală și subpersonală. *Simt* durere; nu creierul meu. *Văd* lucruri; nu ochii mei. Vorbind despre durere, de exemplu, am subliniat:

> O analiză a modului nostru actual de a vorbi despre durere arată că în creier nu se pot descoperi evenimente sau procese care să prezinte caracteristicile presupuselor „fenomene mentale” ale durerii, deoarece modul de a vorbi despre durere este în esență non-mecanic, iar evenimentele și procesele creierului sunt în esență mecanice.
>
> *{Conținut și conștiință,* p. 91)

În ciuda cât de multe avem în comun, Bennett și Hacker îmi denigrează profund munca. Cum pot explica acest lucru? Așa cum se întâmplă adesea în filosofie, este util să ai pe cineva care afirmă, direct și clar, ceea ce alții doar sugerează sau presupun tacit. Bennett și Hacker reușesc să articuleze poziții pe care le combat *indirect* de patruzeci de ani, dar pe care nu le-am putut ataca niciodată direct, din lipsa unui susținător deschis. La fel ca Jerry Fodor, pe care m-am bazat ani de zile pentru a articula cu forță *punctele pe* care doream să le neg - scutindu-mă de a ataca un „om de paie” - Bennett și Hacker îmi oferă o doctrină clară pe

care o pot critica deschis. Sarcina de a-mi organiza gândurile despre aceste probleme ca răspuns la afirmațiile lor a fost foarte revelatoare pentru mine și am încredere că va fi și pentru alții.

CONTEXTUL FILOSOFIC

În această secțiune voi discuta doar despre Hacker, excluzându-l pe coautorul său, Bennett , deoarece punctele pe care le voi critica reprezintă în mod clar contribuția lui Hacker . El repetă, adesea textual, afirmații pe care le face în cartea sa *Wittgenstein: Meaning and Mind* (Blackwell, 1990), afirmații strict filozofice.

Când Hacker mă atacă, iar și iar, pentru că nu am înțeles eroarea mereologică, e ca și cum ai încerca să o înveți pe bunica ta cum să sugă ouă. Cunosc argumentul , deoarece am fost prima care l-a invocat. M-am rătăcit poate când am plecat de la Oxford? Printre filosofii care au preluat distincția mea dintre nivelurile personal și subpersonal, cel puțin unul (Jennifer Hornsby) a conjecturat că este posibil să fi abandonat această cale în lucrările mele ulterioare. Chiar am întors spatele acestei idei bune? Nu. Ar fi cel mai potrivit aici să citez critica mea din 1980 la adresa apărării intuiției de către Searle în spatele cărții „Camera chinezească”:

> Răspunsul sistemelor indică, în opinia mea, destul de corect, că Searle a confundat diferitele niveluri de explicație (și atribuire). *Eu* înțeleg engleza, nu creierul meu - și nici, mai precis, partea corespunzătoare a creierului meu (dacă poate fi măcar izolată) care „procesează” propozițiile percepute și traduce intențiile mele comunicative în acte de vorbire.
>
> *{Științe comportamentale și ale creierului* [1980], 3, p. 429) [8]

(De altfel, Searle a respins sumar această afirmație a mea în răspunsul său la BBS. Sunt curios să văd ce va spune despre distincția dintre nivelurile personal și subpersonal în baza erorii sale reologice .) [9]

Textul canonic pe care Hacker își bazează convingerea despre eroarea mereologică este o singură frază din Sfântul Ludovic:

> În rezumat: Numai despre o ființă umană și despre ceea ce seamănă cu o ființă umană (se comportă ca atare) se poate spune: are senzații; vede, este orb; aude, este surd; este conștient sau inconștient.
>
> *{Investigații filozofice,* paragraful 281)*

Tocmai aici eu și Hacker avem opinii diferite. Îmi face plăcere să-l citez pe Wittgenstein; de fapt, consider că *extind* poziția lui Wittgenstein: văd că roboții și computerele care joacă șah, și da, creierul și părțile sale, *într-adevăr* „seamănă cu o ființă umană vie (pentru că se comportă ca o ființă umană)", iar această asemănare este suficientă pentru a justifica o utilizare corectă a vocabularului psihologic pentru a caracteriza acest comportament. Hacker nu reușește să înțeleagă acest lucru, iar el și Bennett etichetează toate cazurile de astfel de utilizare drept „incoerente", insistând în repetate rânduri că „nu au sens". Deci, cine are dreptate?

Să ne întoarcem în 1969 și să vedem cum am prezentat subiectul atunci:

> Într-un fel, distincția dintre nivelurile personal și subpersonal nu este deloc nouă. Filosofia minții inițiată de Ryle și Wittgenstein este în mare măsură o analiză a conceptelor pe care le folosim la nivel personal, iar lecția care trebuie învățată din atacurile lui Ryle asupra „ipotezelor paramecanice" și din insistența adesea surprinzătoare a lui Wittgenstein că explicațiile se termină mult mai devreme decât am crezut este că nu trebuie să confundăm nivelurile personal și subpersonal. Cu toate acestea, în unele cazuri, această lecție a fost înțeleasă greșit, la fel ca și lecția **că nivelul personal de explicație este singurul nivel de explicație atunci când obiectul studiului îl reprezintă mințile și acțiunile umane .** Într-un sens important, dar limitat, acest lucru este adevărat, deoarece, așa cum vedem în cazul durerii, abandonarea nivelului

*Traducere spaniolă: *Investigații filosofice,* Barcelona, Altaza, 1999. *(Nota traducătorului)*

personal înseamnă încetarea vorbirii despre durere. Într-un alt sens important, este fals și acesta este adesea uitatul. Recunoașterea faptului că există două niveluri de explicație dă naștere obligației de a le corela, iar **aceasta** este **o sarcină care nu se află în afara domeniului filosofului. [...]** Rămâne întrebarea cum se raportează fiecare porțiune a *discursului* despre durere la impulsurile neuronale sau la discursul despre impulsurile neuronale . Această întrebare, și altele similare privind alte fenomene, necesită răspunsuri detaliate chiar și după ce s-a convenit că există diferite tipuri de explicații, diferite niveluri și diferite categorii.

{Conținut și conștiință, pp. 95-96)

Acest pasaj oferă o idee generală despre sarcina pe care mi-am propus-o în ultimii treizeci și cinci de ani. Secțiunile îngroșate evidențiază principalele puncte de dezacord cu Hacker, deoarece calea mea nu este în niciun caz cea pe care a urmat-o el. El își prezintă motivele , care merită o analiză atentă:

a) Întrebările conceptuale preced întrebările de adevăr și falsitate [...] Rezultă că întrebările conceptuale nu pot fi transferate cercetării și experimentării științifice sau procesului de teoretizare științifică.

(pagina 2)

Putem pune sub semnul întrebării prima afirmație. Nu sunt răspunsurile la aceste întrebări conceptuale nici adevărate, nici false ? Conform lui Hacker, nu.

b) Ceea ce sunt adevărul și falsitatea pentru știință, sunt simțul și nonsensul pentru filosofie.

(p. 6)

Astfel, atunci când filosofii greșesc, produc nonsensuri , niciodată falsități, iar când fac o treabă bună nu ar trebui să spunem că au *dreptate* sau că spun *adevărul,* ci pur și simplu că ceea ce spun are sens. [10] Sunt înclinat să cred că propoziția lui

Hacker (b) este pur și simplu *falsă,* nu un nonsens , dar, oricare ar fi ea, a doua afirmație a lui Hacker din (a), în ciuda „rezultă că", nu rezultă. Chiar dacă - întrebările conceptuale „preced" problemele de adevăr și falsitate , este posibil ca oricine dorește să clarifice care sunt răspunsurile bune să investigheze cu sârguință cercetările științifice relevante. Hacker numește această propoziție naturalism quinean †și o respinge cu un argument irelevant: „nu credem că investigația empirică poate rezolva vreo problemă filosofică , la fel cum nu poate rezolva probleme în matematică" (p. 414). Evident că nu; cercetarea empirică nu le rezolvă *, ci* le *informează* și uneori *le ajustează* sau *le revizuiește,* iar apoi aceste probleme uneori *se dizolvă, iar* alteori pot fi rezolvate prin reflecție filosofică suplimentară.

Insistența lui Hacker că filosofia este o *disciplină a priori* Lipsa de continuitate cu știința empirică este principala sursă a problemelor care afectează acest proiect, după cum vom vedea:

> c) Cum pot fi investigate limitele sensului? Numai prin examinarea utilizării cuvintelor. Absurdițile apar adesea atunci când o expresie este folosită contrar regulilor care îi guvernează utilizarea. Expresia în cauză poate fi o expresie cotidiană, netehnică, caz în care regulile de utilizare a acesteia pot fi deduse din utilizarea sa comună și din explicațiile standard ale sensului său. Sau poate fi un termen tehnic dintr-o disciplină, caz în care regulile de utilizare a acestuia trebuie deduse din introducerea termenului de către teoretician și din explicațiile pe care le oferă cu privire la utilizarea sa stipulată. Ambele tipuri de termeni pot fi utilizate greșit, iar când se întâmplă acest lucru, se produce un nonsens - o combinație de cuvinte care este exclusă din limbaj. Fie nu s-a stipulat nimic despre ce înseamnă termenul în contextul aberant în cauză, fie această formă a cuvintelor este de fapt exclusă de o regulă care specifică că nu există așa ceva (de exemplu, că nu există așa ceva ca „la est de Polul Nord"), că aceasta este o combinație de cuvinte care nu are nicio utilitate.

†Referindu-se la filosoful și matematicianul american Willard V.O. Quine (1908–2000), care a pus sub semnul întrebării existența adevărurilor conceptuale non-empirice. *(Nota traducătorului)*

(pagina 6)

Acest pasaj amintește foarte mult de anii 1960 și vreau să-i reamintesc cititorului unele dintre problemele pe care le conține, probleme pe care le credeam rezolvate de mulți ani, deși la vremea respectivă nu aveam o versiune atât de evidentă pe care să o folosim drept țintă pentru criticile noastre.

> Cum pot fi investigate limitele sensului? Numai prin examinarea utilizării cuvintelor.

Să observăm, în primul rând, că, indiferent de ceea ce ar putea spune orice filosof, examinarea utilizării cuvintelor este o investigație empirică, care produce adesea adevăruri și falsuri banale și este supusă corectării prin observații și obiecții standard. Poate că o vagă apreciere a acestei contradicții flagrante l-a determinat pe Hacker să afirme următoarele în cartea sa din 1990:

> Gramatica este autonomă; nu este responsabilă de propoziții factuale, dar acestea trebuie să o presupună. În acest sens, spre deosebire de regulile mijloace-scopuri ale logicii, este arbitrară. Dar este legată de ceea ce nu este arbitrar. Este modelată de natura umană și de natura lumii din jurul nostru.
>
> (p. 148)

Fie ca gramatica să fie autonomă, orice ar însemna asta. Cu toate acestea, nu o poți studia fără să pui întrebări și, chiar dacă îți pui *întrebările,* trebuie să iei în considerare ceea ce spui. Convingerea că această metodă de consultare a propriilor intuiții (gramaticale sau de altă natură) este complet distinctă de cercetarea empirică are o lungă tradiție (care datează nu doar de la Oxford în anii 1960, ci și de la Socrate), dar nu rezistă analizei reflexive.

Acest lucru devine perfect clar dacă comparăm acest stil de filosofie cu antropologia, o cercetare evident empirică ce poate fi făcută bine sau prost. Dacă cineva alege informatori de mâna a doua sau nu reușește să le stăpânească mai întâi

limba, va produce o lucrare de mâna a treia. Din acest motiv, unii antropologi preferă să practice o formă sau alta de *autoantropologie,* în care se folosește pe sine ca informator, poate cu complicitatea unor colegi apropiați ca interlocutori. Natura empirică a demersului este exact aceeași. [Lingviștii] , după cum se știe, practică o formă de antropologie de acest fel și, până acum, știu multe despre capcanele și riscurile încercărilor lor particulare de a descifra intuițiile gramaticale referitoare la limbile lor materne . Este bine cunoscut, de exemplu, cât de dificil este să nu-ți contaminezi propriile intuiții despre gramaticalitate cu idei teoretice alese de tine. De fapt, unii lingviști au ajuns la concluzia că lingviștii teoreticieni nu sunt sau nu ar trebui să fie recrutați ca informatori, deoarece judecățile lor nu sunt inocente . Hacker și alți filozofi cu aceleași idei se confruntă aici cu o provocare semnificativă: cum anume își disting cercetările de autoantropologie, o investigație empirică ce folosește în mod evident aceleași metode și duce la același tip de judecăți? [12]

Oricine crede că filosofii au găsit o metodă de investigație *gramaticală* care este cumva imună (sau independentă, sau „precedentă") la problemele pe care le poate pune o astfel de investigație antropologică ne datorează o explicație detaliată a modului în care aceasta se realizează. Simplele afirmații că *asta fac filosofii* nu fac decât să evite întrebarea. Colegul meu, Avner Baz, îmi amintește de o sugestie interesantă a lui Stanley Cavell^{[13}] pentru a aborda această dificultate: Cavell susține că observațiile filosofului despre ceea ce spunem sunt mai apropiate de *judecățile estetice.* În cuvintele lui Baz, „cineva își prezintă judecata ca fiind *exemplară,* vorbește *în numele* unei comunități" (comunicare personală), iar acest lucru funcționează până la un punct. Dar, din moment ce antropologul încearcă, de asemenea, să găsească cea *mai bună* și *mai coerentă* interpretare a datelor disponibile (ideea lui Quine despre principiul carității și a mea cu privire la - presupunerea raționalității atitudinii intenționale), acest element normativ sau sancționare este deja prezent (deși între paranteze) în cercetarea antropologului. Antropologul nu poate interpreta ceea ce *spun informatorii săi* fără a dezvălui ce *ar trebui să spună aceștia* (în comunitatea lor) în multe situații. Totuși, ceea ce este în mod deliberat omis din demersul antropologului și ceea ce este necesar să

fie apărat în cel al filosofului este *o justificare* pentru următoarea afirmație: Asta fac și spun acești oameni *și cu toții ar trebui să facem la fel.* După cum vom vedea, Hacker nu reușește să identifice comunitatea în numele căreia vorbește, iar acest lucru îi subminează proiectul.

Să revenim aici):

> Prostiile apar adesea atunci când o expresie este folosită împotriva regulilor care îi guvernează utilizarea.

Acest tip de subterfugiu filosofic ar fi trebuit să fie scos în afara legii de mult timp . După cum se știe, Ryle a susținut că a identificat „ erori categorice" apelând la „logica" afirmațiilor de existență, dar să fim serioși: aceasta a fost o cacealma. Îi lipsea o logică articulată a termenilor de existență care să-și susțină afirmațiile. În ciuda popularității acestui tip de discurs, de la Ryle și Wittgenstein până la o mulțime de imitatori, niciun filosof nu a formulat *vreodată* „regulile" pentru utilizarea vreunei expresii comune. Filosofii au formulat sute de judecăți despre utilizări deviante, dar a afirma că „nu am spune-o așa" nu este același lucru cu a enunța o regulă. Lingviștii folosesc un asterisc sau o stea pentru a face același tip de observație și au generat mii de propoziții marcate cu asterisc, cum ar fi - următoarea:

* O ghindă a crescut ca orice stejar.
* Casa era infestată cu șobolani.

Însă, așa cum ne-ar spune orice lingvist, atragerea atenției asupra unei judecăți de utilizare deviantă - chiar și atunci când aceasta indică un model semnificativ și bine descris de deviere - nu este același lucru cu descoperirea regulilor care guvernează astfel de cazuri. Lingviștii lucrează cu sârguință de peste patruzeci de ani pentru a formula regulile sintaxei și semanticii limbajului și există domenii mici în care pot pretinde în mod convingător că au descoperit „regulile".

Dar au întâlnit și mari zone de vagitate. Ce am spune despre această propoziție?

* Pisica a coborât din copac [exemplul lui Jackendoff]

Este aceasta o prostie care contravine „regulilor" verbului **a trepari*?* E greu de spus și e posibil ca uzul să se schimbe. Exemplele de acest fel abundă. Lingviștii au învățat că ceva poate suna destul de ciudat, poate emana un sentiment negativ, fără a încălca prin aceasta vreo regulă clară pe care nimeni nu a reușit să o formuleze și să o apere . Iar ideea *regulilor inefabile* este prea obscurantistă pentru a merita discutată. Intuițiile filosofilor, oricât de bine puse la punct, nu sunt o sursă superioară de dovezi în această investigație evident empirică.

Să revenim la a). Hackerul împarte apoi lexiconul în două:

> Expresia în cauză poate fi o expresie comună, netehnică, caz în care regulile de utilizare a acesteia pot fi deduse din utilizarea sa obișnuită și din explicațiile standard ale semnificației sale. Sau poate fi un termen tehnic dintr-o disciplină, caz în care regulile de utilizare a acestuia trebuie deduse din introducerea termenului de către teoretician și din explicațiile pe care le oferă cu privire la utilizarea sa stipulată.

Sunt tentat să afirm că Hacker pur și simplu greșește (dar nu se contrazice) atunci când sugerează că semnul distinctiv al unui termen tehnic este faptul că este „introdus" de un teoretician care „stipulează" utilizarea sa. Fie spune asta, fie definește „termenul tehnic" atât de restrâns încât mulți termeni pe care în mod normal i-am considera tehnici nu ar fi așa pentru Hacker - iar „termenul tehnic" este un termen tehnic a cărui utilizare Hacker o stipulează aici și acum. Să-l lăsăm, așadar, cu definiția sa a termenilor tehnici, oricât de restrânsă ar fi. Niciunul dintre termenii pe care îi atacă în carte nu este tehnic în acest sens; prin urmare, trebuie să fie termeni „obișnuiți, netehnici" sau hibrizi, o posibilitate pe care Hacker o ia în considerare pe scurt și apoi o respinge în cartea sa din 1990.

> Dacă neurofiziologii, psihologii, experții în inteligență artificială sau filosofii doresc să modifice gramatica existentă pentru a introduce noi moduri de vorbire, pot face acest lucru; dar trebuie să explice noile lor stipulări și să definească condițiile de aplicare . Ceea ce nu se poate face este să se argumenteze că, din moment ce știm ce înseamnă „a gândi", „a vedea" sau „a deduce" și știm ce înseamnă „creier", atunci trebuie să știm ce înseamnă „creierul gândește , vede și deduce". Știm ce înseamnă aceste verbe doar în măsura în care stăpânim utilizarea lor existentă, care nu autorizează aplicarea lor la corp sau la părțile sale, decât într-un mod derivat. Nici noua utilizare „tehnică" nu poate fi amestecată cu cea veche, așa cum fac adesea neurocercetătorii, de exemplu, atunci când teoretizează. Motivul este că acest lucru creează un conflict de reguli și, în consecință, o inconsecvență în utilizarea acestor termeni *de către neurocercetători* .
>
> (pp. 148-149)

Această ultimă afirmație, care se află și în centrul cărții din 2003, este o provocare. Dacă Hacker ar putea să *ne arate* regulile și să ne spună exact cum noile utilizări intră în conflict cu ele, am putea fi în măsură să fim de acord sau nu cu el, dar el doar speculează. Nu are nicio idee care sunt „regulile" pentru utilizarea acestor termeni psihologici de zi cu zi. Și mai revelator este faptul că insistența sa încăpățânată asupra unei metodologii a priori îl împiedică să vadă ce face în acest moment. Să acceptăm că este corectă convingerea sa că posedă o metodă *a priori* care îi oferă º intuiție „prealabilă" a semnificațiilor termenilor săi psihologici actuali . El este totuși obligat să arate cum prolegomenul sau decorul său îl salvează de capcana a ceea ce am putea numi miopie conceptuală: tratarea propriilor *concepte* (posibil restricționate și nefondate) ca fiind obligatorii pentru alții cu scopuri diferite și formațiuni distincte. Cum stabilește că el și cei a căror muncă o critică vorbesc aceeași limbă? Aceasta este, fără îndoială, o întrebare empirică , iar incapacitatea lui Hacker de a o aborda cu suficientă atenție l-a dus pe o cale greșită. Ceea ce face el, de fapt, nu este o filosofie bună, ci o antropologie proastă. S-a orientat către științele cognitive pentru a „examina utilizarea cuvintelor" și nu a

realizat că el însuși își ducea *propriul limbaj cotidian pe un teritoriu* nefamiliar și că intuițiile *sale* nu erau neapărat aplicabile. Când numește utilizarea *lor* „aberantă", face o greșeală de începător .

Utilizarea predicatelor psihologice în teoretizarea de către oamenii de știință cognitivi este, fără îndoială, jargon englezesc, destul de diferit de modul în care vorbesc profesorii de la Oxford și are propriile „reguli ". De unde știu asta? Pentru că am făcut studiul antropologic. (Trebuie să fii naturalist quinean ca să nu faci greșeli atât de simple.) Există un pasaj revelator în care Hacker recunoaște această posibilitate, dar își arată incapacitatea de a o lua în serios:

> Este o descoperire nouă faptul că și creierul desfășoară astfel de activități umane? Sau este o inovație lingvistică introdusă de neurocercetători, psihologi și specialiști în științe cognitive, extinzând astfel utilizarea comună a acestor expresii psihologice din motive teoretice solide? Sau, mai alarmant, este o confuzie conceptuală?
>
> (pp. 70-71)

Hacker optează pentru a treia posibilitate, fără a oferi argumente, în timp ce eu înclin spre primele două împreună. *Există* aici un element de descoperire. Este un fapt empiric și surprinzător că creierul nostru - mai precis, *părți* ale creierului nostru - desfășoară procese *remarcabil de* similare cu conjecturarea, luarea de decizii , credința, tragerea concluziilor și așa mai departe. Și sunt *suficient de similare* cu aceste comportamente la nivel personal pentru a justifica extinderea limbajului obișnuit pentru a le cuprinde. Dacă cineva nu a studiat excelenta muncă științifică pe care această adoptare a atitudinii intenționale a făcut-o posibilă, ar putea crede că este o nebunie să vorbești așa. Dar nu este.

De fapt, asta m-a determinat să-mi dezvolt concepția despre atitudinea intențională. Când am început să petrec timp discutând cu cercetători în informatică și neuroștiințe cognitive, ceea ce m-a impresionat a fost faptul că, fără nicio afectare, fără să le împingă coatele sau să ridice din sprâncene, vorbeau despre cum computerele (și programele, subrutinele și părțile creierului etc.) *vor, gândesc ,*

concluzionează, decid și așa mai departe. Care sunt regulile?, mă întrebam. Și răspunsul la care am ajuns sunt regulile pentru adoptarea atitudinii intenționale. Întrebarea este: Oamenii care lucrează în aceste domenii vorbesc așa și oare atitudinea intențională surprinde cel puțin o parte esențială a „regulilor" care guvernează acest mod de a vorbi? Iar răspunsul (factual) este: Da. [Există], de asemenea, presupun, o întrebare politică : Au dreptul să vorbească așa? Ei bine, faptul că o fac merită cu prisosință, deoarece le permite să genereze ipoteze verificabile, să dezvolte teorii, să analizeze fenomene extrem de complexe în părțile lor cele mai ușor de înțeles etc.

Hacker descoperă și această utilizare omniprezentă a termenilor intenționali în neuroștiințe și este uimit – credeți-mă, uimit! Atât de mulți oameni fac erori conceptuale atât de mari. Și jumătate dintre ei nu știu nimic despre asta. Nu este vorba doar de neurocercetători; sunt informaticieni (și nu doar cei care lucrează în domeniul inteligenței artificiale), etologi cognitivi, biologi celulari, biologi evoluționiști – toți participând cu plăcere la jocul de carte, învățându-și elevii să gândească și să vorbească în acest fel, o pandemie lingvistică. Dacă ați întreba inginerul electrician obișnuit cum funcționează jumătate din aparatele electrocasnice din casa dumneavoastră, răspunsul ar fi plin de termeni intenționali care constituie eroarea mereologică (dacă este într-adevăr o eroare).

Nu este o eroare. Nu atribuim unor părți ale creierului o credință completă (sau o decizie sau o dorință - sau o durere, Dumnezeu știe -) - *aceasta* ar *fi* o eroare. Nu; atribuim acestor părți o formă atenuată de credință și dorință, o credință și o dorință lipsite de multe dintre conotațiile lor cotidiene (referitoare la responsabilitate și înțelegere, de exemplu). Așa cum un copil mic ar putea crede *ceva de genul* că tatăl său este doctor (fără a avea o înțelegere completă a ceea ce este un tată sau un doctor), un [robot] - sau o anumită parte a creierului unei persoane - ar putea crede *ceva de genul* că există o ușă deschisă la câțiva metri în față sau că se întâmplă ceva acolo, în dreapta și așa mai departe. Ani de zile am susținut astfel de utilizări ale intenționalității pentru a caracteriza sisteme complexe, de la computere care joacă șah la termostate, și pentru a caracteriza subsistemele creierului la mai multe

niveluri. Ideea este că, atunci când proiectăm un sistem complex (sau descifrăm designul unui sistem biologic, cum ar fi o persoană sau creierul unei persoane), putem face progrese prin descompunerea persoanei minunate în ansamblu în anumite tipuri de sub-persoane, sisteme agentive care au *o parte* din capacitatea unei persoane, iar apoi acești homunculi pot fi descompuși în agenți și mai simpli, chiar mai puțini oameni și așa mai departe: o regresie *finită,* nu infinită, care atinge punctul culminant atunci când ajungem la agenți atât de stupizi încât pot fi înlocuiți de o mașină. Este posibil ca toate încercările mele de a justifica și explica această mișcare să fie înțelese greșit, dar, din moment ce Bennett și Hacker nu le iau niciodată în considerare, nu sunt în măsură să le evalueze.

Spun asta într-un articol pe care Hacker îl citează de mai multe ori (deși nu și în acest pasaj):

> Am putea fi tentați să ne întrebăm: Sunt componentele subpersonale *sisteme intenționale autentice?* În ce punct al declinului capacității, pe măsură ce coborâm la nivelul neuronilor individuali, *dispare intenționalitatea autentică?* Aceasta nu este o întrebare care să ne fie pusă. Motivele pentru care considerăm un neuron individual (sau un termostat) drept un sistem intențional sunt neglijabile, dar nu zero, iar certitudinea atribuirilor noastre intenționale la cele mai înalte niveluri nu depinde de identificarea unui nivel inferior de intenționalitate autentică.
>
> („Autoportret”, în Guttenplan, catalogat de H și B ca „Dennett, Daniel C. Dennett”, 1994)

Eroarea homunculusului , prin atribuirea *întregii* minți unei părți potrivite a sistemului, nu face decât să amâne analiza și, în consecință, ar genera o regresie infinită, deoarece fiecare postulat nu ar produce niciun progres. Departe de a fi o greșeală , atribuirea unei semi-semi-proto-cvasi-pseudo-intenționalități părților mereologice ale oamenilor este tocmai mișcarea care ne permite să vedem cât de minunați și completi se obțin oameni din părți mecanice stupide. Acesta este ceva

diabolic de greu de imaginat, iar licența poetică pe care o permite atitudinea intențională facilitează substanțial sarcina. [17] Astfel, din poziția mea, Hacker este comic de naiv, precum gramaticianul de modă veche care mustră pe cineva pentru că folosește forma colocvială, dialectală engleză **ain't** și insistă că *nu se spune așa* oamenilor care știu ce vor să spună atunci când o spun. Hacker prevăzuse această perspectivă în cartea sa din 1990 și a descris-o foarte bine:

> Dacă toate acestea [știința cognitivă] trebuie luate în serios , par să demonstreze, în primul rând, că observația gramaticală conform căreia aceste predicate, în sensul lor literal, sunt limitate la ființele umane și la ceea ce se comportă ca ființe umane, este fie pur și simplu greșită, fie prezintă o „inerție semantică" care a fost depășită de progresul științei, deoarece mașinile se comportă într-adevăr ca ființele umane. În al doilea rând, dacă are sens literal să atribuim predicate epistemice și chiar perceptive mașinilor concepute pentru a simula anumite funcții umane și a îndeplini anumite sarcini umane, pare plauzibil să presupunem că creierul uman trebuie să aibă o structură funcțională abstractă similară cu cea a designului mașinii . În acest caz, nu există nicio îndoială că atribuirea diversității predicatelor psihologice creierului uman trebuie să aibă sens până la urmă.
>
> (pp. 160-161)

Exact. Aceasta este teza. Cum o respinge? Nu o face. El spune: „Problemele filozofice apar din confuzia conceptuală. Nu sunt rezolvate prin descoperiri empirice, iar schimbarea conceptuală nu le poate răspunde, ci doar le ascunde sub covor " (p. 161). Întrucât problemele filozofice ale lui Hacker devin depășite, presupun că am putea pur și simplu să le ascundem sub covor, dar aș prefera să le îngropăm cum se cuvine.

DETALII NEUROȘTIINȚIFICE

Când Bennett și Hacker examinează literatura neurologică, criticile lor suferă de o lipsă de varietate. Ei îi citează pe Crick, Edelman, Damasio, Gregory și mulți alții care spun lucruri surprinzător de „incoerente", deoarece acești oameni de știință cad în așa-numita eroare mereologică:

> Expresiile psihologice pe care le folosesc, departe de a fi noi utilizări omonime, sunt invocate în sensul lor obișnuit; altfel, neurocercetătorii nu ar deduce din ele ceea ce deduc ei. Când Crick afirmă că „ceea ce vedem nu este *de fapt* ceea ce este acolo; este ceea ce creierul nostru *crede* că este acolo", este important să presupunem că „crede" are conotațiile sale normale - că nu înseamnă același lucru cu un anumit termen *nou* , „crede*". Deoarece face parte din teoria lui Crick că credința este rezultatul unei *interpretări* bazate pe *experiențe* sau *informații anterioare* (și nu rezultatul unei interpretări* bazate pe experiențe* și informații* anterioare).
>
> (p. 75)

De fapt, Bennett și Hacker pur și simplu greșesc (dar nu cad în prostii). Crick intenționează ca întreaga sa explicație (în acest caz, una destul de banală și nicidecum controversată) să fie înțeleasă la nivel subpersonal. Interpretarea în cauză *nu este* a *experienței* (la nivel personal), ci, de exemplu, a *datelor din fluxul ventral,* iar procesul de interpretare este evident presupus a fi un proces subpersonal. Un alt pasaj în aceeași ordine de idei:

> În mod similar, atunci când [JZ] Young afirmă că creierul conține cunoștințe și informații, care sunt codificate în creier „în același mod în care cunoștințele sunt înregistrate în cărți sau computere ", el se referă la cunoștințe, nu la cunoștințe ‡- deoarece cunoștințele și informațiile, nu cunoștințele* și

‡*A cunoaște:* un neologism introdus de Chomsky pentru a descrie relația dintre un vorbitor și regulile limbii sale și, mai general, dintre o persoană și cunoașterea sa, denotând un tip de cunoaștere în mare măsură implicită sau tacită. Vezi Chomsky, N.,

informațiile *, pot fi înregistrate în cărți și computere.

Autorii nu au făcut absolut nimic pentru a demonstra că nu există niciun concept de cunoaștere sau informație care să poată fi codificat atât în cărți, cât și în creier. În științele cognitive , există o literatură vastă și complexă despre conceptul de informație - și despre cel de cunoaștere (este suficient să ne gândim la discuțiile lui Chomsky despre „cogniție"* ca răspuns la o critică care a fost foarte asemănătoare cu cea a autorilor) - iar ignorarea de către autori a acestor dezbateri anterioare arată că nu își iau sarcina foarte în serios. Ar putea fi citate multe alte exemple de același fel. Autorii noștri au o idee, eroarea mereologică, și o folosesc sistematic, fără a lua în considerare detaliile câtuși de puțin. De fiecare dată, citează pasajul nelegitim - și ar fi putut găsi mult mai multe exemple de atribuire a atitudinilor intenționale subsistemelor cerebrale - și apoi pur și simplu îl declară o contradicție pentru că cade în eroarea lor. Nici măcar o dată nu încearcă să demonstreze cum autorul în cauză, comitând o eroare atât de teribilă, este condus inevitabil la o greșeală sau contradicție reală. Cine și-ar fi putut imagina că filosofia neuroștiinței este atât de ușoară?

Să luăm în considerare discuția lor despre subiectul fascinant și controversat al imaginilor mentale. În primul rând, ei demonstrează - deși mă îndoiesc că cineva a contestat vreodată acest lucru - că imaginația creativă și imaginile mentale sunt de fapt fenomene destul de distincte și independente. Apoi vine lovitura de grație: „O zonă senzorială aranjată topografic nu este o imagine a nimic; nu există imagini în creier, iar creierul nu *are* imagini " (p. 183). Ca cineva care a susținut cu seriozitate ani de zile că nu ar trebui să ne grăbim să concluzionăm că „ imaginile mentale" implică imagini reale în creier și că tracturile retinotopice situate acolo ar putea să nu *funcționeze* ca imagini în procesarea creierului, trebuie să subliniez că afirmația slabă a lui Bennett și Hacker nu este de niciun ajutor. Pur și simplu este irelevant

Knowledge of Language, New York, Praeger, 1986 (traducere în spaniolă: *El conocimiento del lenguaje,* Madrid, Alianza, 1989). *(Nota traducătorului)*

dacă „spunem” că creierul are *imagini.* Întrebarea dacă unele dintre fasciculele de stimuli ale creierului care posedă în mod manifest proprietățile geometrice ale imaginilor *funcționează de fapt* ca imagini este o întrebare empirică și unul al cărui răspuns este aproape. Analiza filosofică nu poate rezolva problema – decât printr-o insistență profund reacționară că aceste structuri de date iconice, din care informațiile sunt aparent extrase în același mod în care oamenii (la nivel personal) extrag vizual informații din imagini publice, nu sunt *considerate* imagini. Astfel de accese de ofuscare creează o problemă serioasă de credibilitate pentru filosofie în ochii științei cognitive .

De fapt, există probleme conceptuale serioase în modul în care specialiștii în știința cognitivă vorbesc despre imagini, cunoștințe, reprezentări, informații și orice altceva. Dar este o sarcină dificilă și laborioasă să *demonstrăm* că terminologia utilizată este utilizată în moduri care îi derutează serios pe teoreticieni. Adevărul este că, în cea mai mare parte, acești termeni, așa cum se găsesc în știința cognitivă, sunt de fapt doar „limbaj obișnuit” - nu termeni tehnici [explicit] definiți în vreo teorie. Adesea, teoreticienii au considerat util să vorbească, într-un mod *oarecum* impresionist, despre informațiile procesate , deciziile luate și reprezentările consultate. Și atunci când sunt provocați, în loc să facă ceea ce ar face un filosof - adică să-și *definească termenii mai precis* - indică spre modelele lor și spun: „Uite: iată mecanismele în acțiune, efectuând procesarea informațiilor despre care vorbeam”. Și modelele funcționează. *Se comportă* așa cum trebuie să se comporte pentru a face să apară acel homunculus anume, astfel încât nu mai este nevoie de alte obiecții cu privire la ceea ce a fost atribuit exact sistemului.

Există însă și multe ocazii când entuziasmul teoreticienilor pentru interpretările intenționate ale modelelor lor îi duce în eroare.19 [De] exemplu, în dezbaterea despre imagistică, s-au făcut greșeli sub forma unor interpretări exagerate - de exemplu, de către Stephen Kosslyn - care trebuie corectate. Nu este vorba că a vorbi despre hărți sau imagini în neuroștiințe este complet *deplasat ,* ci mai degrabă că trebuie făcut cu mare atenție , ceea ce nu este întotdeauna cazul. Poate ajuta filosofia? Da, spun Bennett și Hacker: „Poate explica - așa cum am

explicat - de ce imaginile mentale nu sunt ca fotografiile eterice și de ce nu pot fi rotite în spațiul mental" (p. 405). Această abordare maximalistă nu este utilă. Nu poți stabili ce se întâmplă de fapt în creier atunci când oamenii sunt cufundați în imagini mentale spunând că nivelul personal nu este nivelul subpersonal. Teoreticienii știu deja acest lucru; nu fac *această* greșeală. De fapt, sunt gânditori foarte atenți și subtili, iar unii dintre ei *încă* vor să vorbească despre imagini din creier care funcționează ca niște imagini. S-ar putea să aibă dreptate. [20] Filosofi precum Hacker își pot pierde interesul față de subiect odată ce acesta este situat la nivel subpersonal, [21] dar atunci nu ar trebui să facă greșeala de a critica un domeniu despre care știu atât de puțin.

Uneori, autorii sunt atât de naivi și încrezători încât efectul este comic, ca în cazul condamnării severe a lui David Marr:

> A vedea nu înseamnă a descoperi ceva dintr-o imagine sau dintr-un fascicul de lumină care cade pe retină. Pentru că nu este posibil să descoperi nimic, în acest sens , din ceva ce nu putem percepe (noi nu percepem fasciculul de lumină care cade pe retinele noastre, ceea ce percepem este orice fasciculul de lumină ne permite să percepem).
>
> (p. 144)

Am înțeles. Din moment ce nu *percepem* fasciculul de lumină care ne atinge retinele, e evident că nu *descoperim* nimic. Marr nu era idiot. A înțeles asta. Deci, ce zici de teoria lui Marr despre *procesele subpersonale* ale vederii?

> În plus, *nu este deloc clar* [sublinierea mea] cum faptul că mintea are acces la presupusele *descrieri neuronale* îi va permite persoanei să *vadă.* Și dacă Marr insistă (pe bună dreptate) că persoana, nu mintea, este cea care vede, cum se explică trecerea de la prezența unei descrieri a unui model tridimensional codificat în creier la experiența de a vedea ceea ce este în fața ochilor? *Aceasta* nu este, de fapt, o problemă empirică pe care cercetările ulterioare ar trebui să o rezolve. Este produsul unei confuzii conceptuale și ceea ce are nevoie este o clarificare.

(p. 147)

Dimpotrivă, aș argumenta că este o problemă filozofică ce se rezolvă făcându-i pe cei care cred că *„nu este deloc clară ” să înțeleagă* că teoria lui Marr *poate* explica repertoriul de abilități asociate cu vederea. Marr a presupus, [mai mult] sau mai puțin, că cititorii săi pot înțelege cum un model al creierului care oferă o reprezentare tridimensională a lumii ar putea explica cum o creatură cu exact acest tip de creier ar putea vedea, dar dacă unii cititori nu au reușit să înțeleagă, un filosof ar fi probabil specialistul potrivit pentru a le-o explica. Simpla *afirmație* că Marr suferă de o confuzie conceptuală are, așa cum a spus-o atât de bine Russell , toate avantajele furtului față de munca cinstită.

> Pentru că a vedea ceva este exersarea unei capacități, o utilizare a facultății vizuale — *nu* [accentul meu] procesarea informațiilor în sens semantic și nici producerea unei descrieri în creier.
>
> (p. 147)

Acesta *nu este* un alt caz de furt. Ceea ce trebuie explicat este capacitatea „facultății vizuale”, iar această capacitate este explicată în termenii capacităților combinate ale părților sale, printre ale căror activități se numără crearea și consultarea descrierilor (dacă pot fi numite așa). Aceste exemple ar putea fi multiplicate ad nauseam:

> Nu are sens, decât ca o figură de stil echivocă, să spunem, așa cum face Le Doux, că este „posibil ca creierul să știe că ceva este bun sau rău înainte de a ști exact ce este”.
>
> (p. 152)

Dar cine este confuz? Nu LeDoux, nici cititorii lui LeDoux, dacă citesc cu atenție, *deoarece* își vor da seama că a găsit o modalitate foarte bună de a formula teza surprinzătoare conform căreia un circuit specializat din creier poate discerne

ceva ca fiind, de exemplu, periculos sau dezirabil, pe baza unui fel de triaj rapid *care §*se încheie *înainte ca* informațiile să treacă la acele rețele care completează identificarea stimulului. (Da, da, știu. Doar o persoană - un medic, o asistentă medicală sau cineva de genul acesta - poate pune în aplicare comportamentul pe care îl numim *triaj* - vorbesc în sens metonimic. Cititorul ar trebui să se obișnuiască cu el.)

Pe scurt, ceea ce le spun colegilor mei din neuroștiințe este că nu există niciun argument de apărat aici. Autorii susțin că aproape toți cei implicați în neuroștiința cognitivă comit o gafă conceptuală destul de simplă. Eu spun că toate acuzațiile ar trebui respinse până când autorii prezintă dovezi demne de luat în considerare. Oferă autorii altceva care ar putea avea vreo valoare pentru neuroștiințe? Nu oferă nicio teorie sau model pozitiv, nici vreo sugestie despre cum ar putea fi construite astfel de teorii sau modele, deoarece, evident , acesta nu ar fi domeniul filosofiei. Interpretările lor „corecte" ale comisurotomiei și ale vederii oarbe, de exemplu, constau în reprezentări banale ale fenomenelor descrise și deloc în explicații. Au dreptate într-o oarecare măsură: așa apar aceste fenomene remarcabile. Acum, cum ar trebui să le explicăm? Explicația trebuie să se oprească undeva, așa cum spunea Wittgenstein, dar nu aici. Bennett și Hacker citează cu consternare unele dintre remarcile extrem de disprețuitoare despre filosofie făcute de Glynn, Crick, Edelman, Zeki și alții (pp. 396–398). Având în vedere forța acestui argument, este de înțeles că neurologii sunt atât de puțin interesați. [23]

§Un sistem medical pentru selectarea și clasificarea pacienților în situații de urgență sau dezastre, în funcție de șansele lor de supraviețuire, nevoile terapeutice și resursele disponibile. *(Nota traducătorului)*

PENTRU A ÎNLOCI CONȘTIINȚA ÎN CREIER

Răspuns la Fundamentele filozofice ale neuroștiinței, *de Bennett și Hacker*

IOAN SEARLE

Aceasta este o carte lungă, de peste 450 de pagini, care acoperă un număr vast de probleme. Conține numeroase obiecții la adresa ideilor mele, precum și o anexă dedicată în mod special criticii mele. Îmi voi limita observațiile la anumite probleme centrale din carte și voi răspunde la ceea ce consider a fi cele mai importante critici ale lui Bennett și Hacker. Dar nu voi încerca să discut toate problemele substanțiale ridicate în cartea lor.

Întrucât majoritatea comentariilor mele vor fi critice, vreau să încep prin a nota câteva puncte importante de acord. Autorii au dreptate când subliniază că, în percepție, percepem de obicei obiecte reale din lume și nu reproduceri sau imagini interne ale obiectelor . De asemenea, au dreptate când observă că relația normală pe care o avem cu propriile experiențe nu este epistemică. Este o greșeală să ne gândim la relația noastră cu percepțiile noastre pe baza modelului „accesului privilegiat" sau a oricărui alt tip de „acces" epistemic, iar modelul „introspecției ", prin care *ne spectamus intro* (ne uităm în interior) la propria minte, este iremediabil de confuz. Am subliniat toate aceste [aspecte] în diverse scrieri și mă bucur să văd că suntem de acord asupra acestor aspecte.

Le-am trimis prima variantă a acestui articol și au avut amabilitatea să mi-au semnalat pasajele în care, în opinia lor, le-am interpretat sau formulat greșit ideile, ceea ce mi-a permis să corectez anumite neînțelegeri.

Concepția sa fundamentală despre fenomenele mentale este, în opinia mea, eronată și voi încerca să explic de ce mai jos. Pentru a evidenția diferența dintre ideile mele și ale lui, voi începe cu un scurt rezumat al unora dintre concepțiile mele și apoi le voi expune clar pe ale lui prin contrast. Pentru concizie, îmi voi limita discuția la conștiință, deși este posibil să se facă, *mutatis mutandis,* Observații similare asupra intenționalității .

CONȘTIINȚA CA FENOMEN BIOLOGIC

1. Prin definiție, conștiința constă în stări (voi folosi „stări" pentru a desemna stări, procese, evenimente etc.) care sunt *calitative și subiective.* Lăsând la o parte patologiile, stările conștiente apar doar ca parte a unui *singur câmp unificat al conștiinței.* Conștiința este calitativă în sensul că pentru fiecare stare conștientă există un anumit caracter calitativ, un aspect al cum-este-sau-cum-se-simte. De exemplu, caracterul calitativ al consumului de bere este diferit de caracterul - calitativ al ascultării *Simfoniei a IX-a* de Beethoven . *Aceste stări sunt subiective* în sensul ontologic că există doar în măsura în care sunt experimentate de un subiect uman sau animal. Și sunt *unificate* în sensul că orice stare conștientă, cum ar fi senzația prezentă a tastaturii sub degetele mele, există ca parte a unei stări conștiente mai ample, câmpul meu prezent al conștiinței. Având în vedere caracterul lor subiectiv și calitativ , aceste stări sunt uneori numite *„qualia".* În general, nu consider că este un concept util, deoarece implică o distincție între acele stări conștiente care sunt calitative și cele care nu sunt, iar din punctul meu de vedere, nu există o astfel de distincție. „Conștiință" și *„qualia"* sunt pur și simplu termeni cu aceeași arie de aplicare. Totuși , întrucât Bennett și Hacker neagă existența *qualiilor ,* în acest text voi folosi termenul pentru a evidenția

punctele de dezacord. Când spun că există stări conștiente, mă refer la faptul că există *qualii* . Când spun că există stări conștiente, ei înseamnă cu totul altceva, după cum vom vedea.

Fenomenele conștiente sunt fenomene concrete care se produc în spațiu-timp. Nu sunt entități abstracte precum numerele. Anterior credeam că calitativitatea, subiectivitatea și unitatea sunt trei caracteristici distinctive ale conștiinței, dar după ce am reflectat asupra acestui lucru, mi se pare clar că fiecare dintre ele o implică pe următoarea . Sunt trei aspecte diferite ale esenței conștiinței: subiectivitatea calitativă și subiectivitatea unificată.

2. Aceste stări, *qualia,* sunt cauzate în întregime de procese cerebrale. Nu suntem complet siguri care sunt mecanismele cauzale, dar se pare că excitația neuronilor la nivelul sinapselor joacă un rol funcțional deosebit de important .

3. Stările conștiente există în creier. Acestea sunt realizate în creier ca trăsături de nivel înalt sau ca un sistem în ansamblu. De exemplu, gândurile conștiente despre bunicile noastre sunt procese care au loc în creier, dar, din câte știm, niciun neuron singular nu poate provoca sau efectua gânduri despre o bunică. Conștiința este o caracteristică a creierului la un nivel superior celui al neuronilor individuali. [2]

Evident, sunt multe de spus despre conștiință, iar unele dintre ele le-am spus și în altă parte. Consider principiile enunțate în punctele 1, 2 și 3 ca fiind adevăruri ale unui bun simț științific mai mult sau mai puțin cultivat. Dar interesul discuției actuale constă în faptul că, în mod incredibil, *Bennett și Hacker neagă toate cele trei principii.*

Ei susțin că au demonstrat că noțiunea de *qualia,* noțiunea de caracter calitativ al experiențelor conștiente, este „incoerentă ". De asemenea, spun că: „ne deosebim de Searle [...] atunci când spune că fenomenele mentale sunt cauzate de procese neurofiziologice ale creierului și sunt ele însele caracteristici ale creierului" (p. 446).

Chiar și pentru a înțelege magnitudinea provocării sale și consecințele acesteia pentru filosofie și neuroștiințe , haideți să aplicăm cele trei principii ale mele unui

exemplu din viața reală. Chiar acum, văd o mână în fața mea. Care sunt părțile componente ale evenimentului în care văd mâna? Ei bine, în primul rând, trebuie să existe o mână acolo și trebuie să existe un anumit impact asupra aparatului meu vizual și neurobiologic (îl voi scuti pe cititor de detalii). În cazul normal, cum ar fi cel al vederii non-orbe , acest impact va produce o experiență vizuală conștientă, în sensul meu de eveniment calitativ și subiectiv, un *quale.* Vreau să subliniez că experiența vizuală are toate caracteristicile pe care tocmai le-am menționat: este calitativă, subiectivă și există ca parte a unui câmp unificat. Este cauzată de procesele cerebrale și există în creier. Astfel, avem trei componente ale scenei vizuale: cel care percepe, obiectul perceput și experiența vizuală calitativă. O mare parte din cele mai bune lucrări din neuroștiințe reprezintă un efort de a explica modul în care procesele cerebrale cauzează experiența vizuală și unde și cum are loc această experiență în creier. În mod uimitor, Bennett și Hacker neagă existența experienței vizuale în acest sens, în sensul quale . *Este* corect să subliniem, așa cum fac și ei, că ceea ce percep este o mână și nu o experiență vizuală, dar să negăm categoric existența unei experiențe vizuale calitative, în sensul quale . Percepția vizuală este, ca să spun cel puțin, bizară. Ce se întâmplă când închid ochii, de exemplu ? Experiența vizuală calitativă se termină. Acesta este motivul pentru care nu mai văd mâna, pentru că nu mai am experiența vizuală. Rețineți că prezența mâinii este esențială pentru ca eu să pot vedea efectiv mâna, dar nu este esențială pentru existența experienței vizuale, deoarece, în cazul halucinațiilor, pot avea o experiență indistinguizabilă fără ca vreo mână să fie prezentă.

Bennett și Hacker nu sunt primii autori care neagă existența qualiilor , dar negarea lor nu este motivată de materialismul lipsit de sens al celor care se tem că, dacă acceptă existența unor fenomene mentale ireductibil subiective, se vor trezi cu Descartes ca tovarăș de pat. Așadar, ce motivează, ce *ar putea* motiva, o negare a existenței stărilor conștiente, așa cum le-am definit?

VIZIUNEA WITTGENSTEINIANĂ

Cea mai bună modalitate de a înțelege cartea lui Bennett și Hacker este să o vedem ca pe o aplicație a filosofiei minții a lui Wittgenstein la neuroștiința contemporană. O mare parte din originalitatea cărții constă în faptul că acest lucru nu a mai fost făcut până acum. Poziția lui Bennett și Hacker este , din câte știu, unică în dezbaterile contemporane despre filosofia minții.

Care este, așadar, interpretarea lor asupra filosofiei minții a lui Wittgenstein și cum o aplică neuroștiințelor? Un pasaj cheie din Wittgenstein, pe care îl citează, este:

„Numai despre ființele umane și despre ceea ce seamănă cu o ființă umană (se comportă ca atare) se poate spune: are senzații; vede, este orb; aude, este surd; este conștient sau inconștient" (p. 71).

Cu alte cuvinte, *are sens doar* să atribuim predicate mentale unui lucru care este o ființă umană sau se comportă ca atare. Și care este rolul comportamentului în atribuirea acestor predicate? Comportamentul oferă nu numai baza pentru inducerea prezenței fenomenelor mentale, ci și criteriile logice. Ar trebui să înțelegem manifestarea comportamentului ca fiind *criteriul logic* pentru aplicarea acestor concepte. Există o legătură semantică între comportamentul extern și conceptele mentale, deoarece numai o ființă capabilă să prezinte o anumită formă de comportament poate fi considerată posedă un fenomen mental. Două expresii cheie sunt următoarele (p. 83): *„Criteriile pentru atribuirea unui predicat psihologic sunt parțial constitutive ale semnificației acelui predicat"* și, pe aceeași pagină: *„Creierul nu îndeplinește criteriile pentru a fi un posibil subiect al predicatelor psihologice."*

Acum, din moment ce Bennett și Hacker acceptă această concepție wittgensteiniană, ei cred că consecința sa logică imediată este că conștiința nu poate exista în creier și că creierul nu poate desfășura activități mentale, cum ar fi gândirea sau percepția, deoarece este incapabil să manifeste comportamentul adecvat (p. 83). Numai despre persoana în ansamblu sau, în cazul animalelor, despre animalul întreg, putem spune că suferă sau este furios, deoarece numai

animalul întreg este capabil să manifeste comportamentul care constituie o parte din condițiile de aplicare a conceptului în cauză. Având în vedere legătura logică dintre stările mentale și comportament, nu putem face separarea tradițională între fenomenul mental și manifestarea sa externă. Această conexiune logică explică de ce putem vedea literalmente că cineva este furios sau suferă, sau este conștient sau inconștient. Mai mult, Bennett și Hacker sunt acum obligați să nege existența - qualiilor , deoarece dacă acestea ar exista, ar exista în creier, iar acest lucru contrazice teza conform căreia conștiința nu poate exista în creier. Pentru ei, ideea că conștiința există în creier nu este doar falsă: este lipsită de sens. Este ca și cum ai spune că există conștiință în numere prime.

De asemenea, fac numeroase mișcări wittgensteiniene cu privire la modul în care copiii învață vocabularul mental și de ce avem un vocabular mental. Din câte îmi amintesc, nu folosesc explicit ideea de „joc lingvistic", dar este implicită în întreaga carte. Jocul lingvistic pe care îl jucăm cu cuvinte mentale necesită criterii comportamentale observabile pentru aplicarea sa .

Aceasta este punctul lor de vedere și cred că majoritatea tezelor de fond ale cărții derivă de fapt din ea. Este o bază validă pentru a critica studiile contemporane? Este suficientă pentru a respinge ideea că conștiința constă dintr-o subiectivitate unificată, calitativă, cauzată de procesele cerebrale și realizată în creier? Nu cred, din motivele care urmează. Să presupunem că au dreptate în ceea ce privește baza logică, necesitatea unui comportament observabil public pentru a juca jocul limbajului , imposibilitatea limbajului privat și tot restul. Ce urmează? Niciuna dintre concluziile lor spectaculoase . Odată ce acceptăm, așa cum trebuie și așa cum a făcut Wittgenstein, că există o distincție între durere și - comportamentul dureros, între sentimentele de furie și comportamentul furios, între gândire și comportamentul de gândire și așa mai departe, atunci putem concentra atenția cercetărilor noastre neurobiologice asupra durerii, sentimentelor de furie, experienței vizuale și așa mai departe și să uităm de comportament. Așa cum vechii behavioristi au confundat dovezile comportamentale pentru stările

mentale cu existența stărilor mentale în sine, adepții lui Wittgenstein comit o - eroare mai subtilă, dar fundamental similară, atunci când confundă criteriile de aplicare a conceptelor mentale cu stările mentale în sine. Adică, ei confundă criteriile comportamentale pentru *atribuirea* predicatelor psihologice cu *faptele atribuite de aceste* predicate psihologice, ceea ce este o greșeală gravă.

Să presupunem că Wittgenstein are dreptate când spune că nu am putea avea un vocabular legat de durere decât dacă ar exista forme de comportament dureros care ar putea fi exprimate public. În mod similar, dacă mă întreb: „Ce fapt despre mine face ca eu să simt durere?", nu pot găsi niciun fapt despre comportamentul meu care să facă din asta să simt durere. Faptul despre mine care face ca eu să simt durere este faptul că am un anumit tip de senzație neplăcută. Și ceea ce este valabil pentru durere este valabil și pentru furie, gând și orice altceva. Chiar dacă argumentul lui Wittgenstein este sută la sută corect ca analiză filosofică a modului în care funcționează vocabularul , *putem totuși* deduce întotdeauna, în orice caz individual, existența unui sentiment interior, calitativ și subiectiv din manifestarea sa în comportamentul exterior. Bennett și Hacker subliniază că, deși distincția se poate face în cazuri individuale, nu este posibil să nu existe niciodată manifestări public observabile ale durerii, deoarece altfel nu am putea folosi vocabularul legat de durere. Să presupunem că au dreptate în această privință. În orice caz, atunci când investigăm ontologia durerii – nu condițiile pentru a juca jocul limbajului, ci ontologia fenomenului în sine – putem ignora comportamentul extern *și* ne putem concentra pur și simplu pe a descoperi cum creierul provoacă senzații interne.

Rețineți că în pasajul citat, Wittgenstein vorbește despre ceea ce putem *spune.*

„Numai despre ființele umane și despre ceea ce seamănă cu o ființă umană (se comportă ca atare) se poate *spune: «* are senzații; vede, este orb; aude, este surd; este conștient sau inconștient»" (p. 71; sublinierea îmi aparține).

Dar să presupunem că eliminăm cuvântul „a spune" din acest pasaj și îl scriem astfel:

„Numai ființele umane și lucrurile care se comportă ca ființe umane pot *fi cu adevărat* conștiente." Dacă luăm aceasta ca pe o afirmație conceptuală sau logică,

este evident falsă. Să presupunem că o luăm ca:

„Este un adevăr conceptual sau logic că doar lucrurile care se comportă ca ființele umane pot fi cu adevărat conștiente."

Însă acest lucru, ca adevăr conceptual sau logic, pare a fi pur și simplu greșit. De exemplu, moluștele și crustaceele, cum ar fi stridiile și crabii, nu se comportă deloc „ca ființele umane ", dar acest fapt în sine nu rezolvă întrebarea dacă sunt sau nu conștiente. Indiferent de cât de mult diferă comportamentul lor de comportamentul uman, stridiile ar fi totuși conștiente dacă ar avea tipul potrivit de procese neurologice în sistemul lor nervos. Să presupunem că avem o știință perfectă a creierului și știm exact cum apare conștiința la oameni și la animalele superioare. Dacă am descoperi atunci că mecanismul care produce conștiința este prezent la stridii, dar nu la melci, am avea motive întemeiate - ce-i drept, motive copleșitoare - să presupunem că stridiile sunt conștiente, iar melcii probabil nu. Însăși existența conștiinței nu are nicio legătură cu comportamentul, deși , la ființele umane, existența comportamentului este (un criteriu) esențial pentru funcționarea jocului limbajului. Întrebarea „care dintre animalele inferioare sunt conștiente?" nu poate fi rezolvată prin analiză lingvistică.

Wittgenstein oferă o explicație generală a funcționării vocabularului mental. El subliniază că este o greșeală să interpretăm funcționarea jocului lingvistic din modelul dovezilor inductive externe ale prezenței fenomenelor private interne. „Un proces intern necesită un criteriu extern", ne amintește el. Dar chiar dacă acceptăm această explicație a vocabularului, nimic nu ne împiedică să oferim o versiune neurobiologică a modului în care stările de conștiință sunt cauzate de procesele cerebrale și se desfășoară în sistemele cerebrale. Mai mult, cerința ca sistemul, persoana în ansamblu, să fie capabilă să manifeste comportament nu implică faptul că nu poate exista un element al sistemului, creierul, care să fie locul proceselor conștiente. Acesta este un alt subiect, pe care îl voi explica mai detaliat în secțiunea următoare.

Eroarea, pe scurt, constă în confuzia regulilor de utilizare a cuvintelor cu ontologia. Așa cum behaviorismul vechi a confundat *testele* stărilor mentale cu

ontologia stărilor mentale, acest behaviorism logic wittgensteinian confundă *baza atribuirii* cu *faptul atribuit.* Este o eroare să spunem că condițiile pentru - funcționarea corectă a jocului limbajului sunt condiții pentru existența fenomenelor în cauză. Să presupunem că avem o știință perfectă a creierului, astfel încât să știm cum produce creierul durerea. Să presupunem că construim o mașină care este capabilă de conștiință și chiar capabilă să fie conștientă de senzația de durere. Am putea proiecta o mașină astfel încât să nu prezinte niciun fel de comportament dureros. Ar depinde de noi. Există, fără îndoială, forme de boli reale în care oamenii experimentează durere fără a prezenta un comportament dureros. În unele cazuri de sindrom Guillain-Barré, pacientul este pe deplin conștient, dar complet paralizat, total incapabil să prezinte un comportament compatibil cu starea sa mentală. Bennett și Hacker subliniază că acest lucru nu ar putea fi cazul pentru toată durerea - că nimeni care suferă de durere nu ar prezenta vreodată un comportament dureros - pentru că atunci nu am putea aplica cuvinte. Chiar dacă acest lucru este corect, este o condiție pentru buna funcționare a limbajului, nu o condiție pentru existența durerii.

REZUMATUL ARGUMENTELOR PREZENTATE PÂNĂ ACUM

Cred că odată ce această eroare fundamentală este eliminată, argumentul central al cărții se destramă. Voi discuta argumentele detaliate ale autorilor noștri într-o clipă, dar mai întâi vreau să rezum argumentele prezentate până acum. Wittgenstein susține că o condiție pentru posibilitatea unui limbaj de a descrie - fenomenele mentale interne este manifestările observabile public ale acelor fenomene. Comportamentul nu este doar o dovadă inductivă, ci un criteriu pentru aplicarea conceptelor. Să presupunem, de dragul argumentului, că are dreptate . Bennett și Hacker cred că rezultă că fenomenele mentale nu ar putea exista în creier, deoarece creierul nu poate manifesta comportamentul care servește drept criteriu. Dar acest lucru nu rezultă din ideea că un comportament observabil este

un criteriu pentru aplicarea limbajului la mental. Tot ce ar putea rezulta din aceasta este că, dacă putem vorbi despre stări mentale în creier, atunci creierul trebuie să facă parte dintr-un mecanism cauzal capabil să producă comportament (voi aborda acest punct mai detaliat în secțiunea următoare). Și de obicei este. Dar chiar și în cazurile în care nu este, trebuie să distingem însăși existența fenomenelor mentale de posibilitatea de a vorbi despre ele. Faptul care face ca eu să am durere atunci când de fapt am durere este existența unui anumit tip de senzație. Indiferent dacă manifest sau nu acea senzație în comportament, este irelevant pentru existența ei .

ARGUMENTUL PRINCIPAL AL CĂRȚII: EROAREA MEREOLOGICĂ

Cel mai important argument al cărții, cel pe care îl repetă iar și iar, este demontarea a ceea ce Bennett și Hacker numesc „eroarea mereologică", pe care o definesc ca atribuirea părților a ceea ce are sens doar atunci când este atribuit întregului. Forma tipică pe care o ia această eroare, potrivit lor, este aceea că neurocercetătorii spun lucruri precum că *creierul* gândește, percepe, are încredere, pune întrebări, decide etc., când, de fapt, caracterizarea corectă ar fi că *întreaga persoană* gândește, percepe, are încredere, pune întrebări, decide etc. Eroarea constă în atribuirea părții, creierului, a ceea ce are sens doar atunci când este atribuit întregului, persoanei în totalitatea sa. Sper că este clar că acest lucru rezultă din viziunea wittgensteiniană : deoarece partea, adică creierul, nu poate manifesta un comportament conștient și deoarece comportamentul conștient este esențial pentru atribuirea conștiinței, nu putem atribui durere creierului.

Voi face o observație pedantică, a cărei importanță va deveni clară mai târziu. Conform explicației autorilor noștri, acesta nu poate fi, strict vorbind, un caz de eroare mereologică - eroarea de a atribui părții ceea ce are sens doar atunci când este atribuit întregului - pentru că, dacă ar fi așa, am putea elimina eroarea pur și simplu adăugând o referință la restul corpului. Relația creierului cu restul corpului

este , fără îndoială, o relație parte-întreg. Creierul este o parte a corpului meu. Bennett și Hacker spun că numai *persoana* poate fi subiectul atribuirilor psihologice, nu doar creierul. Dar persoana nu este legată de creier în felul în care întregul este legat de parte. Aceasta nu implică faptul că persoana este ceva distinct de corp sau că este „deasupra" acestuia. Din păcate, autorii nu ne spun niciodată ce este o persoană și totuși cred că acest lucru este crucial pentru expunerea în ansamblu , ba chiar pentru întreaga discuție. Ceea ce ei numesc eroare mereologică este descrisă mai precis ca o greșeală de categorie în sensul lui Ryle. Conform lui Bennett și Hacker, oamenii aparțin unei categorii logice diferite față de creier și, din acest motiv, atribuirile psihologice oamenilor nu au sens atunci când sunt atribuite creierului. Voi reveni asupra acestui punct mai târziu.

După cum știu bine autorii noștri, există de fapt (cel puțin) trei tipuri diferite de atribuiri subpersonale ale fenomenelor psihologice , iar argumentele pe care le folosesc împotriva uneia dintre ele nu sunt neapărat aplicabile celorlalte. În primul rând, există creierul ca subiect și agent (de exemplu, „Creierul gândește"). În al doilea rând, există creierul ca locație a proceselor psihologice (de exemplu, „Gândirea are loc în creier"). Și în al treilea rând, există microelementele ca agenți (de exemplu , „Neuronii individuali gândesc"). Să examinăm fiecare dintre acestea în ordine.

În primul rând, creierul ca subiect și agent. După cum am menționat anterior, este destul de comun atât în literatura filosofică, cât și în cea neurobiologică să se descrie cunoașterea folosind verbe cognitive al căror subiect este „creierul". Astfel, este obișnuit să spunem că creierul percepe , creierul gândește, creierul decide etc., lucru pe care Bennett și Hacker îl consideră inacceptabil din motivul pe care l-am explicat: că creierul nu poate manifesta comportamentul adecvat. În vorbirea de zi cu zi, trebuie să spunem că persoana decide. Eu sunt cel care decide să voteze pentru candidatul democrat, nu creierul meu.

Argumentul pentru această concluzie, așa cum am menționat anterior, este că, din moment ce creierul nu poate manifesta comportament, nu poate fi subiectul unor predicate psihologice. Dar odată ce ne dăm seama cât de slab este acest

argument, ne putem gândi la vreun alt motiv pentru a respinge atribuirea proceselor psihologice creierului? Sunt de acord cu autorii noștri că este ciudat să spunem, de exemplu, „Creierul meu a decis să voteze cu democrații la ultimele alegeri". De ce? Voi reveni asupra acestui punct mai târziu.

În al doilea rând, creierul ca loc. O a doua formă de atribuire, complet diferită de prima, este de a preciza *unde* au loc procesele și evenimentele psihologice. Și aici teza este că acestea au loc în creier. Bennett și Hacker sunt conștienți de distincția dintre tratarea creierului ca agent și subiect al proceselor psihologice și tratarea lui ca loc al proceselor psihologice ; dar se opun ambelor idei. Ei cred atât că creierul nu poate gândi, cât și că gândirea nu poate avea loc în creier. Dar pentru a demonstra că creierul nu poate fi locul unor astfel de procese, ar avea nevoie de un argument independent, un argument pe care nu l-am putut găsi. Strict vorbind, chiar dacă argumentul wittgensteinian ar fi valid, nu ar putea fi folosit împotriva tuturor acestor atribuiri. De ce nu? Argumentul afirmă că agentul unui proces psihologic trebuie să fie un sistem capabil să prezinte comportamentul adecvat. Astfel, în exemplul nostru de vedere, un sistem
A fi capabil să vadă trebuie să fie capabil să manifeste comportamentul adecvat. Prin urmare, nu putem spune despre un creier că vede; putem spune acest lucru doar despre întregul sistem, adică despre persoană. Dar acest lucru nu ne împiedică să identificăm experiența vizuală drept o componentă a vederii și nici să localizăm experiența vizuală în creier. Tot ce impune argumentul wittgensteinian este ca creierul să facă parte dintr-un mecanism cauzal în cadrul unui sistem global capabil să producă comportament . Și această condiție poate fi îndeplinită chiar și atunci când anumite procese psihologice sunt localizate în creier.

Pentru a înțelege acest lucru, să luăm în considerare o analogie. Să presupunem că cineva spune: „Nu ar trebui să spunem că stomacul și restul tractului digestiv digeră alimentele. Doar întreaga persoană poate digera alimentele." Într-un fel, acest lucru este adevărat. Dar rețineți că, în scopul cercetării modului în care funcționează digestia, ne putem întreba unde și cum au loc procesele digestive specifice. Și răspunsul este că acestea au loc în stomac și

în restul tractului digestiv. Acum, în același mod, cineva ar putea insista că eu, persoana, sunt cea care percepe și gândește conștient, și nu creierul meu. În mod similar, cineva s-ar putea întreba atunci în ce parte a anatomiei au loc procesele conștiente, iar răspunsul evident este că acestea au loc în creier. Bennett și Hacker sunt conștienți de distincție, dar nu par să realizeze că afirmația că creierul nu poate fi locul proceselor psihologice ar necesita un argument diferit, iar eu nu găsesc un astfel de argument în cartea lor. Ei spun: „Localizarea *evenimentului în care o persoană gândește un anumit gând* este locul în care se află persoana atunci când gândul îi apare" (p. 180). Cu siguranță este adevărat, dar asta nu înseamnă că gândurile mele nu apar și în capul meu. Anumite gânduri îmi vin chiar în acest moment. Unde? În această cameră. Unde anume, în această cameră? În creierul meu. De fapt, odată cu dezvoltarea imagisticii prin rezonanță magnetică funcțională (fMRI) și a altor tehnici de imagistică , ne apropiem din ce în ce mai mult de a putea spune exact unde în creier au loc gândurile.

LOCALIZAREA STĂRILOR CONȘTIENTE ȘI A proceselor cerebrale ca și cauză a acestor stări

Concepția pe care am anticipat-o este că procesele mentale conștiente au loc în creier și sunt cauzate de procese neuronale de nivel inferior. Care este concepția lui Bennett și Hacker ? În măsura în care neagă că procesele mentale au loc în creier, cred că sunt incapabili să ofere o explicație coerentă a localizării și cauzei conștiinței. Ei cred că procesele neuronale sunt o condiție necesară a conștiinței, dar nu enunță și, în opinia mea, nu pot enunța teza evidentă că, în circumstanțele potrivite, procesele neurobiologice sunt cauza suficientă a conștiinței. Stările mele conștiente actuale, toate fiind qualia , sunt cauzate de procese neuronale de nivel inferior din creier.

Cred că unele dintre cele mai slabe argumente din carte se referă la această problemă a localizării și cauzei stărilor conștiente . Autorii noștri afirmă: „Nu

există *așa ceva* ca un proces mental (cum ar fi recitarea alfabetului în imaginație) care să aibă loc într-o parte a animalului, fie că este vorba de rinichi sau de creier" (p. 112). „Ceea ce se întâmplă *în* creier sunt procese neuronale, care trebuie să aibă loc pentru ca persoana, posesoarea creierului, să treacă *prin* procese mentale relevante" (p. 112).

Cred că acest pasaj conține o eroare profundă, pe care o voi examina pas cu pas.

Să presupunem că recit alfabetul în tăcere, „în mintea mea", cum am putea spune. Acesta este un eveniment real, din lumea reală. Ca toate evenimentele reale, are loc în spațiu-timp . Deci, unde s-a întâmplat? Bennett și Hacker spun că eu, persoana respectivă, *am trecut prin* procesul mental. Acest lucru este, fără îndoială, adevărat, dar unde anume în spațiu-timp a avut loc evenimentul mental conștient, situat temporal și spațial, al recitării mele în tăcere a alfabetului? Autorii noștri nu pot răspunde la această întrebare decât, poate, spunând că s-a întâmplat în New York, în această cameră sau ceva de genul acesta. Dar asta nu este suficient. Cred că este evident că un eveniment conștient, un set de *qualia,* a avut loc în creierul meu. Și, așa cum am mai spus, odată cu avansul tehnicilor de imagistică, cum ar fi RMN-ul funcțional, suntem mult mai aproape de a putea spune exact unde s-a întâmplat .

Negarea realității și a localizării spațiale a *qualiilor* îi împiedică pe autorii noștri să ofere o explicație *cauzală coerentă* a relației dintre procesele neuronale și evenimentele mentale. Ei spun că procesele neuronale sunt cauzal *necesare* („trebuie să se întâmple") pentru ca un proces mental să aibă loc. Dar, în acest context , trebuie să știm ce este cauzal *suficient,* ce m-a determinat să „trec printr-un" proces mental. Și ori de câte ori vorbim despre cauze, trebuie să spunem exact ce anume ce cauzează. În explicația mea, evenimentul mental conștient este cauzat în întregime de creier și are loc în creier. În acest context, excitația anumitor neuroni este cauzal suficientă pentru a produce acele *qualii,* acele evenimente mentale conștiente . Ce spun ei? Nu pot spune că excitația neuronilor a cauzat

qualiile, experiența calitativă, pentru că au negat existența *qualiilor.* Nu pot spune că excitația neuronilor a cauzat comportamentul extern, pentru că nu a existat niciun comportament extern. Deci, care este exact natura procesului mental prin care am trecut și ce anume l-a determinat să se producă? Neuronii s-au activat ... și apoi ce? Neuronii trebuie să activeze pentru ca eu să „trec prin" procese mentale. Dar în ce constă trecerea prin procese mentale dacă nu există *qualia* sau comportament? Bennett și Hacker nu au un răspuns la aceste întrebări și, luând în considerare teoria lor în ansamblu, nu cred că o vor avea vreodată.

Din perspectiva wittgensteiniană, dacă creierul nu poate manifesta - comportament, atunci nu poate fi nici subiectul, nici agentul atribuțiilor mentale. Cred că aceasta este o greșeală. Dar, indiferent dacă este sau nu, este necesar să se facă distincția între argumentul că creierul nu poate fi *subiectul* verbelor psihologice și argumentul că creierul nu poate fi *locul* proceselor psihologice. Și Bennett și Hacker nu reușesc să distingă între aceste două argumente . Ei sunt conștienți de distincția dintre creier ca subiect și creier ca loc și neagă ambele: că creierul poate fi subiectul și că poate fi locul. Dar nu oferă un argument independent împotriva afirmației că creierul este locul, să zicem, al proceselor de gândire. Să presupunem că suntem de acord că sună ciudat să spunem: „Creierul meu gândește...". Chiar și așa, pot exista în continuare procese de gândire active în creierul meu atunci când *gândesc.* Cel mai mult ce poate stabili argumentul wittgensteinian este că nu ar trebui să ne gândim la creier ca la un subiect sau agent. Dar nu rezultă din aceasta că nu este *locul* proceselor corespunzătoare. Argumentul său împotriva *creierului ca subiect* nu se extinde la *creier ca loc.*

Întrebarea „Unde au loc procesele conștiente?" nu este mai abstrusă din punct de vedere filosofic decât întrebarea „Unde are loc procesul digestiv?". Procesele cognitive sunt procese biologice reale în aceeași măsură în care este digestia. Iar răspunsul la ambele întrebări este evident. Digestia are loc în stomac și în restul tractului digestiv; conștiința are loc în creier și poate în alte părți ale sistemului nervos central.

ATRIBUIRI ALE STĂRILOR PSIHOLOGICE: METAFORICE ȘI LITERAL, RELATIV LA OBSERVATOR ȘI INDEPENDENT DE OBSERVATOR

În al treilea rând, neuronii ca subiecți și agenți. O a treia formă de atribuire la care se opun Bennett și Hacker este atribuirea proceselor psihologice unor părți ale creierului. Astfel, de exemplu, ei îl citează pe Blakemore când spune că neuronii percep, decid, fac inferențe etc. Ei cred că acesta este, de asemenea, un exemplu de simplă eroare logică.

Înțeleg că, atunci când este interpretată corect, aceasta nu este nimic mai mult decât o metaforă inofensivă, sau cel puțin poate fi. Într-adevăr, literatura științifică face atribuiri similare stomacului. Se spune că stomacul știe când sunt necesare anumite substanțe chimice pentru a digera anumite tipuri de carbohidrați care ajung la el . În opinia mea, aceste metafore sunt, sau pot fi, inofensive, cu condiția să distingem clar între sensul literal, independent de observator, în care deduc sau primesc informații, și simțurile metaforice, relative la observator, în care spunem că neuronii noștri fac anumite inferențe sau percep anumite fenomene. Este mai ușor să facem greșeala de a confunda simțurile reale, independente de observator, cu simțurile metaforice, relative la observator, atunci când avem de-a face cu creierul decât cu alte organe, din motivul evident că procesele psihologice intrinseci, independente de observator, au loc în creier și nu în stomac sau în restul tractului digestiv. Sunt convins că Bennett și Hacker au dreptate să creadă că unii dintre autorii pe care îi critică nu înțeleg clar distincția dintre atribuțiile independente de observator și cele legate de observator ale acestor fenomene. Dintr-o perspectivă semantică, principalul vinovat este noțiunea de „informație”. O confuzie suplimentară apare din utilizarea de către ingineri a termenului „informație” în sensul teoriei informației, care nu are nicio legătură cu informația în sensul de a avea informații că se întâmplă cutare sau cutare lucru. Ni se spune, de exemplu, că creierul procesează informații. Ei bine, într-un fel, este adevărat. Obțin informații prin percepție și mă gândesc la ele, apoi deriv informații noi prin inferență. Problema este că în creier (de exemplu, în nucleul geniculat lateral) au

loc tot felul de procese subpersonale, care pot fi descrise *ca* instanțe de gândire despre informații, dar, desigur, nu există informații literale acolo. Există doar excitații neuronale, care sunt traduse în informații conștiente la sfârșitul procesului, dar în sine le lipsește conținut semantic. Bennett și Hacker discută clar distincția dintre sensul de „informație" în teoria informației și sensul intențional. Ei se referă la ele ca sensul „ingineresc" și respectiv sensul „semantic", dar nu găsesc în cartea lor o declarație clară a distincției dintre „informațiile" independente de observator și cele dependente de observator. Eu am informații independente de observator despre numărul meu de telefon. Cartea de telefon conține informații dependente de observator despre același număr de telefon. Nu am nicio reținere în a discuta despre informații și procesarea informațiilor în creier, atâta timp cât aceste distincții sunt clare.

Pentru a rezuma situația actuală, am ridicat trei obiecții principale la -argumentele lui Bennett și Hacker . În primul rând, teza lui Wittgenstein conform căreia comportamentul este un criteriu pentru atribuirea fenomenelor mentale unui limbaj public, chiar dacă este adevărată, nu infirmă concepția că conștiința poate exista în creier. În al doilea rând, odată ce distingem între creier ca subiect și creier ca loc, ne dăm seama că Bennett și Hacker nu au un argument independent -împotriva creierului ca loc. De fapt, din câte știm, este adevărat că locația tuturor proceselor noastre conștiente este în creier. În al treilea rând, atribuirea stărilor psihologice unor entități subpersonale, cum ar fi neuronii, poate fi inofensivă atâta timp cât este clar că este o utilizare metaforică. Dacă distingem între literal și metaforic și între relativul observatorului și independentul observatorului , nu este nimic neapărat greșit în astfel de atribuiri.

ARGUMENTE ÎMPOTRIVA *QUALIEI*

Pentru ca argumentul wittgensteinian să funcționeze, Bennett și Hacker au nevoie de un argument separat împotriva *qualiilor.* De ce? Ei bine, dacă *qualiile* ar exista,

ar trebui să fie localizate undeva, iar cea mai evidentă locație este creierul. Iar acest lucru ar fi incompatibil cu teoria lor în ansamblu. Acum voi examina argumentele lor împotriva *qualiilor.*

Așa cum am mai spus, conștiința constă în procese mentale subiective, calitative și unificate care au loc în creierul uman, în spațiul fizic al craniului, probabil localizat în mare parte în sistemul talamocortical. Teza mea este că durerile, durerile și mâncărimile noastre, de exemplu, sunt *subiective,* în sensul că pot exista doar în măsura în care sunt experimentate de un subiect real; că sunt *calitative* ; și că apar ca *parte a unui câmp conștient unificat.* Bennett și Hacker cred că au argumente împotriva tezei mele. În primul rând , folosind un argument folosit de Ryle, ei susțin că afirmația că orice durere, furnicături sau mâncărime poate fi experimentată doar de un singur subiect nu este nimic mai mult decât o aserțiune gramaticală trivială, fără consecințe ontologice . Ei spun că, exact în același mod, un zâmbet trebuie să fie întotdeauna zâmbetul cuiva, un strănut trebuie să fie strănutul cuiva sau, pentru a folosi un exemplu de la Ryle, o paradă a unui portar într-un meci trebuie să fie parada cuiva. În acest sens, intimitatea strănuturilor, zâmbetelor și mângâierilor nu dovedește nimic semnificativ din punct de vedere ontologic . Același lucru se spune și despre dureri, dureri și mâncărimi. Da, trebuie să fie durerea, furnicătura sau mâncărimea cuiva , dar acesta este doar un aspect gramatical banal, care nu le acordă niciun statut special.

Răspunsul la aceasta, implicit în ceea ce am spus deja, este că expresiile care se referă la stări conștiente nu numai că au caracteristica *gramaticală* de a necesita un sintagm nominal personal pentru identificarea unor cazuri particulare, ci mai degrabă că *subiectivitatea* fenomenelor este legată de însăși natura lor calitativă. Și acesta nu este pur și simplu un aspect gramatical. Într-o durere, o furnicătură sau o mâncărime, există o senzație calitativă specifică , iar aceasta este esențială pentru existența durerilor, furnicăturilor și mâncărimii. Această senzație calitativă face parte din subiectivitatea ontologică în cauză, spre deosebire de caracteristicile , de exemplu, ale paradelor unui portar.

Bennett și Hacker prezintă argumente separate împotriva existenței *qualiilor,*

dar aceste argumente mi se par extrem de slabe. Ei spun că mirosul unui liliac și cel al unui trandafir vor avea același caracter calitativ dacă sunt la fel de plăcute sau neplăcute (pp. 275-276). În mod surprinzător, ei presupun că „calitatea" este o chestiune de grade de plăcere sau neplăcere . Dar nu despre asta este vorba. *Qualiile* mirosului de trandafiri și mirosul de liliac nu sunt constituite de gradul în care sunt plăcute sau neplăcute. Nu este vorba despre asta. Ceea ce se înțelege prin termenul *qualia* este că *natura* experiențelor este diferită . Deși *qualiile* posedă trăsăturile caracteristice ale plăcerii sau neplăcerii, esența care le definește este sensul *calitativ* al experienței. Răspunsul lui Bennett și Hacker la această întrebare mi se pare, de asemenea, inadecvat. Ei susțin că, dacă nu definim *qualiile* ca o chestiune de plăcere sau neplăcere, atunci trebuie să individualizăm experiența pe baza obiectului său: mirosul unui *liliac* sau mirosul unui *trandafir.* Și adaugă că identificarea experienței prin obiectul său nu înseamnă identificarea a nimic subiectiv legat de experiență, deoarece trandafirii au o existență obiectivă. Din nou, mi se pare o greșeală. Desigur, de obicei identificăm caracterul experiențelor noastre perceptive prin cauzele lor, adică prin obiectul intențional care ne face să avem experiența caracteristică. Dar se poate avea experiența și se poate individualiza experiența fără cauze. Dacă se dovedește că actele mele de a mirosi trandafirul și de a mirosi liliacul au fost ambele halucinații, acest lucru nu ar afecta *deloc diferențele dintre qualia.* Aspectele calitative ale mirosului unui trandafir și ale mirosului unui liliac rămân aceleași, indiferent dacă există sau nu un obiect intențional în fiecare caz. Adesea discriminăm mirosurile fără a avea nicio idee ce le-a cauzat . Pe scurt, a spune că atunci când descriem *qualia, o* facem de obicei din perspectiva obiectelor lor - adică a fenomenelor care le provoacă - nu constituie o obiecție la existența qualia cu caracteristici calitative distinctive. Noțiunea de *qualia poate* fi definită în termeni de experiență conștientă pur și simplu prin abstracționarea experienței conștiente de obiectul său intenționat. *Qualia* constau în experiențe conștiente, indiferent de modul în care alegem să le identificăm. Este demn de remarcat faptul că chimiștii care lucrează în industria parfumurilor încearcă să sintetizeze substanțe chimice care reproduc puterile cauzale ale unor

lucruri precum trandafirii și liliacul. Ei încearcă să producă *qualia* de același tip cu cele produse de florile reale.

Bennett și Hacker invocă un al treilea argument care, în opinia mea, este la fel de greșit. Ei spun că, de obicei, persoane diferite pot simți aceeași durere sau pot avea aceeași durere de cap. Dacă eu și cititorul mergem la o petrecere și bem prea mult vin, a doua zi vom avea aceeași durere de cap sau, dacă amândoi suferim de aceeași boală care provoacă dureri de stomac, vom avea aceeași durere. Cred că acest lucru demonstrează încă o dată că nu înțeleg ce încearcă filosofii să spună despre „intimitatea" durerilor . „Intimitatea" poate fi o metaforă slabă, dar este irelevantă aici. Ideea este că atunci când filosofii vorbesc despre aceeași durere, ei se referă la același *tip* , nu la aceeași *instanță.* Ceea ce ne interesează atunci când vorbim despre intimitatea durerilor nu este faptul că persoane diferite nu pot experimenta același tip de durere . Desigur că pot. Ceea ce ne interesează este că, ca exemplu, durerea pe care o experimentăm ei sau eu *există doar în măsura în care este percepută de* un anumit subiect conștient.

LOCALIZAREA DURERILOR

Am spus că toate stările conștiente există în creier. Dar cum rămâne cu durerea la picior? Fără îndoială, este în picior, nu în creier. Bennett și Hacker obiectează la ideile mele în această privință și, întrucât aceste idei sunt la prima vedere contraintuitive , vreau să le clarific. Cred că, dacă vom enunța clar toate faptele, întrebările privind localizarea durerii vor avea răspunsuri evidente.

În primul rând, există spațiu real, adică spațiu fizic. Există un singur spațiu real, fizic, și tot ce se află în el este legat spațial de tot ce există. Astăzi, după Einstein, concepem spațiul și timpul ca un continuum spațiu-timp unic , iar amplasamentele lucrurilor sunt specificate în raport cu un sistem coordonat. În scopurile noastre, spațiul este ordonat logic. Să luăm în considerare tranzitivitatea lui „în": dacă scaunul este în cameră și camera este în casă, atunci scaunul este în casă. Toate evenimentele din lumea reală au loc în spațiu și timp fizic. Unele dintre aceste evenimente nu au limite bine definite - gândiți-vă la Marea Depresiune sau

la Reforma Protestantă, de exemplu - dar, în orice caz, și ca toate celelalte evenimente, ele au loc în spațiu și timp.

Să ne îndreptăm acum atenția către spațiul experimentat și fenomenologic al corpului. Să presupunem că sufăr o rană la picior. Aceasta activează o secvență de excitații neuronale care urcă în coloana vertebrală, trec prin tractul Lissauer dorsolateral, accesează centrii durerii din creierul meu și, ca urmare, *simt durere în picior.* Această descriere este, fără îndoială, corectă, astfel încât, dacă, de exemplu, un medic m-ar întreba unde simt durerea, aș arăta spre picior și nu spre cap. Adică, pentru a indica localizarea durerii, aș arăta spre membrele mele anatomice în spațiul real. Întrebarea este acum: care este exact relația dintre spațiul fizic real și spațiul experimentat și fenomenologic al corpului?

Pentru a răspunde la această întrebare, trebuie să ne întrebăm cum creează creierul spațiul fenomenologic al corpului. Creierul creează o imagine a corpului, o conștientizare fenomenologică a părților corpului, a stării lor și a relațiilor dintre ele. Creierul creează o conștientizare a piciorului meu inclus în imaginea corpului și, în consecință , atunci când simt durere, o conștientizare a durerii așa cum este localizată în piciorul meu. Pe scurt, creierul creează un spațiu corporal fenomenologic real și durere în acel spațiu corporal. Nu există nicio îndoială cu privire la realitatea fenomenologică a acestor fenomene; singura întrebare este: care este exact relația dintre spațiul fenomenologic al corpului și spațiul fizic real al corpului meu?

Problemele apar atunci când încercăm să tratăm spațiul fenomenologic al corpului ca și cum ar fi identic cu spațiul fizic real al corpului. Rețineți că tranzitivitatea lui „în” nu funcționează dacă încercăm să trecem de la spațiul fenomenologic la spațiul fizic. Simt o durere în picior, iar piciorul meu este în cameră, dar durerea nu este în cameră. De ce nu? Puzzle-ul relației dintre spațiul fenomenologic și spațiul fizic devine și mai complicat atunci când luăm în considerare durerea în membrele fantomă. Pacientul simte o durere în picior, dar nu are picior. Durerea este reală, dar unde este? Bennett și Hacker răspund la această întrebare astfel: „Deci, ei simt cu adevărat o durere acolo unde ar trebui să

fie piciorul lor (adică, în membrul lor fantomă)" (p. 125). Dar un membru fantomă nu este un tip de membru, precum un membru rănit sau un membru ars de soare. Membrele fantomă nu există ca obiecte în spațiul real. Așadar, dacă luăm în considerare faptul că observația autorilor noștri se referă literalmente la spațiul fizic , ajungem la rezultatul absurd că pacientul „are o durere acolo unde i-ar fi piciorul", adică în pat. Durerea este exact acolo, sub cearșaf. De ce este absurd acest lucru? Pentru că în spațiul fizic al paturilor și cearșafurilor nu există durere. Durerea poate exista doar în spațiile fenomenologice ale corpului. Și dacă luăm în considerare faptul că pacientul se referă la imaginea fenomenologică a corpului său, ceea ce spune este adevărat. Bărbatul simte o durere în picior și, deși nu are un picior fizic, are totuși un picior fantomă fenomenologic în imaginea corpului său.

Totuși – și acesta este punctul crucial – durerea este un eveniment real, din lumea reală; prin urmare, trebuie să aibă o locație în spațiu-timp real. Nu poate fi în piciorul bărbatului pentru că nu are picior. Nu poate fi acolo unde ar fi piciorul dacă bărbatul ar avea unul, pentru că nu există nimic între cearșafuri. Evident , este în piciorul său fantomă, dar piciorul său fantomă nu este un obiect cu o locație spațială ca parte a corpului, precum un picior real. Am încredere că este evident că, în spațiul fizic al lumii reale, locația spațială a durerii fantomă a piciorului fantomă este în imaginea corpului, care este în creier. În spațiul fizic real, atât durerea reală a piciorului, cât și durerea piciorului fantomă se află în creier, împreună cu restul imaginii corpului.

ESTE O PERSOANĂ UN CREIER (ÎN INTERIORUL UNUI CORP)?

Voi aborda acum problema menționată anterior, aparentul caracter nepotrivit al atribuirii activităților psihologice creierului, mai degrabă decât persoanei în ansamblu. Mi-am exprimat dezacordul cu argumentele pe care le folosesc autorii noștri pentru a demonstra că creierul nu poate conține procese psihologice, dar sunt de acord că sună ciudat să spunem, de exemplu, „Creierul meu a decis să

voteze cu democrații". De ce sună greșit? De ce extindem atât de ușor utilizarea anumitor predicate care se referă la persoană la un anumit aspect al persoanei, cum ar fi corpul acesteia? Luați în considerare următoarele patru propoziții:

1. Cântăresc 75 de kilograme.
2. Pot deosebi albastrul de violet când le văd.
3. Am decis să votez cu democrații.
4. Dețin o proprietate în orașul Berkeley.

În cazul 1, nu am ezita să înlocuim „eu" cu „corpul meu " - adică, *cântăresc* 75 de kilograme dacă și numai dacă *corpul meu* cântărește 75 de kilograme. Ambele afirmații par echivalente. De asemenea, nu văd nicio problemă în a face o substituție similară în propoziția 2. Creierul meu, în special sistemul meu vizual, inclusiv ochii mei, poate distinge albastrul de violet. Dar cred că am ezita mult mai mult să facem o substituție similară în propoziția 3. Dacă vreau să exprim că am decis să votez pentru democrați, pare destul de derutant să spun că creierul meu sau creierul din corpul meu a decis să voteze pentru democrați. O substituție analogă în propoziția 4 este, în opinia mea, și mai ciudată. A spune că creierul din corpul meu deține o proprietate în Berkeley sau că corpul meu deține o proprietate în Berkeley sună în mod evident ciudat. Bennett și Hacker resping substituția din propozițiile 2 și 3. Invocând argumentul wittgensteinian, aceștia nu ar admite nici „Sistemul meu vizual (din creierul meu) poate distinge albastrul de violet", nici „Creierul meu a decis să voteze pentru democrați". Am expus deja motivele pentru respingerea argumentului wittgensteinian; totuși, să admitem că sună impropriu din punct de vedere logic să spunem că creierul meu a decis să voteze pentru democrați. Chiar dacă argumentul wittgensteinian este eronat, trebuie să abordăm această nepotrivire. Pentru mulți neurobiologi și filozofi, li se pare perfect natural să atribuie activități psihologice creierului.

Cum rezolvăm această dispută? Cred că, la nivelul la care dezbatem această problemă acum, există o modalitate destul de ușoară de a rezolva aparentul

conflict. Ori de câte ori avem o formulare alternativă pentru orice afirmație *că S este,* o primă modalitate de a aborda validitatea formulării alternative este să ne întrebăm ce cauzează faptul *că p,* dacă *S* exprimă propoziția *că p.* Nu avem nicio problemă cu prima propoziție, „Cântăresc 75 de kilograme", când o înlocuim cu „Corpul meu cântărește 75 de kilograme", deoarece știm ce fapt despre mine mă face să cântăresc 75 de kilograme - și anume, că asta cântărește corpul meu. Nu am nicio obiecție la o schimbare similară pentru propoziția 2, deoarece dacă ne întrebăm ce fapt despre mine mă face să pot distinge albastrul de violet, faptul în cauză este că sistemul meu vizual este capabil să distingă albastrul de violet. Dar propozițiile 3 și 4 sunt foarte diferite de 1 și 2, deoarece necesită nu numai existența și caracteristicile unui creier integrat într-un corp, ci mai degrabă ca acest creier să fie situat într-un context social și să fie capabil să desfășoare acțiuni sociale . În cazul propoziției 3, spre deosebire de 4, putem face abstracție de situația socială și identifica o componentă pur psihologică . Având în vedere situația mea socială și politică și în acest context, anumite activități care au loc în creierul meu constituie faptul că am decis să votez pentru democrați. Aceeași - ezitare care ne împiedică să atribuim creierului ceva ce necesită ca persoana întrupată în corp să fie situată social ne face, de asemenea, reticenți în a atribui creierului orice fel de acțiune sau statut agentiv. Astfel, deși în creierul meu au loc anumite activități care constituie faptul că am decis să votez pentru democrați, suntem mai reticenți în a atribui creierului astfel de procese raționale de luare a deciziilor decât, de exemplu, abilitățile perceptive . Nu am nicio problemă să spun că „sistemul meu vizual poate distinge roșul de violet", dar sunt mult mai reticent să spun că „sistemul meu talamocortical a decis să voteze pentru democrați". În cazul propoziției 4, „Dețin o proprietate în orașul Berkeley", nu pare să existe nimic ce putem abstrage și atribui anatomiei. Numai în virtutea situației mele sociale și a relațiilor în care mă aflu pot deține o proprietate. Proprietarul unei proprietăți este, în fapt, un creier integrat într-un corp, dar un astfel de creier nu poate deține o proprietate decât în virtutea caracteristicilor sale sociale și juridice și, prin urmare, nu există nicio posibilitate de a face vreo atribuire anatomică sau

psihologică care să fie constitutivă pentru aspectele relevante ale faptelor.

Acestea sunt întrebări filozofice interesante, dar cred că neurobiologii le pot evita perfect. În loc să se preocupe de măsura în care ar trebui să trateze capacitatea de acțiune rațională ca o caracteristică a anatomiei neuronale, ar trebui să se concentreze asupra celui de-al doilea punct, și anume, faptul stabilit că procesele psihologice care constituie capacitatea de acțiune rațională conștientă au loc în creier și pot fi investigate ca atare. În scopul cercetării neurobiologice, este suficient să se studieze creierul ca mecanism cauzal și locație anatomică .

parte a problemei sinelui.3 De ce trebuie să postulăm un sine ca ceva adăugat la secvența experiențelor noastre și la realizarea lor anatomică? Nu pentru că există vreo superanatomie dincolo de anatomie sau superexperiențe dincolo de experiență . Nu există nimic mai mult decât creierul integrat într-un corp și experiențele care au loc în acel creier. Cu toate acestea, așa cum am argumentat, trebuie să postulăm un sine, dar este un postulat pur formal. Nu este o entitate suplimentară, ci un fel de principiu organizator al creierului și al experiențelor sale.

NATURA FILOSOFIEI

În a doua anexă a cărții, Bennett și Hacker critică ideile mele, atât explicația mea asupra problemelor filosofiei minții, cât și abordarea mea generală asupra filosofiei. Ei consideră că concepția mea despre filosofie este greșită în mai multe privințe. Din experiența mea, disputele despre natura filosofiei tind să fie fără rezultat și, de obicei, nu fac altceva decât să exprime preferințe pentru diferite proiecte de cercetare. Am fost educat filosofic la Oxford în anii 1950, unde am fost atât student, cât și lector universitar , iar ortodoxia predominantă atunci era că filosofia se ocupa de limbaj și de utilizarea cuvintelor. Dacă cineva spune că filosofia se ocupă în întregime de limbaj, mi se pare că exprimă o preferință. Este aproximativ echivalent cu: „Prefer munca filosofică care se ocupă de limbaj și intenționez să fac doar muncă filosofică despre limbaj”. Ceea ce am descoperit

este că tehnicile pe care le folosisem pentru a analiza limbajul funcționau cu alte fenomene, și anume fenomenele mentale și ontologia socială . Așadar, metodele pe care le folosesc sunt o continuare a metodelor filosofiei analitice, dar se extind dincolo de domeniul original al filosofiei limbajului și al filosofiei lingvistice în care am fost educat.

Un punct cheie de dezacord între abordarea mea și cea a lui Bennett și Hacker este insistența lor că filosofia nu poate fi teoretică, că nu oferă explicații teoretice generale. Cu toții suntem de acord că filosofia este, într-un fel, în esență conceptuală, dar întrebarea este: care este natura analizei conceptuale și care este rezultatul său final? Din propria mea experiență, pot spune că un rezultat conceptual este important doar ca parte a unei teorii generale . Așadar, dacă mă uit la propria mea istorie intelectuală, am propus o teorie generală a actelor de vorbire și a sensului, o teorie generală a intenționalității, o teorie generală a raționalității și o teorie generală a naturii societății, a ontologiei sociale . Dacă cineva ar spune: „Ei bine, nu poți avea teorii în filosofie", răspunsul meu ar fi: „Ei bine, iată câteva teorii generale". O analiză generală a actului de a promite, de exemplu , capătă mult mai multă putere atunci când este încorporată într-o teorie generală a limbajului și a actelor de vorbire. Analizele filozofice ale acțiunii și percepției devin mult mai profunde atunci când sunt integrate într-o teorie generală a intenționalității, iar același lucru este valabil și în alte cazuri.

Într-un fel, Bennett și Hacker îmi prezintă greșit ideile despre relația dintre filosofie și știință. Ceea ce spun nu este că *toate* problemele filosofice pot fi transformate în probleme științifice printr-o analiză conceptuală atentă. Dimpotrivă, cred că doar un număr mic de probleme filosofice admit o soluție în științele naturii. Problema vieții este una dintre ele și am încredere că problema minte-corp va fi una în viitor. Însă majoritatea problemelor care i-au preocupat pe marii filozofi greci, cum ar fi natura vieții bune, forma unei societăți juste sau cel mai bun sistem de organizare socială, nu sunt, în opinia mea, genul de lucruri care pot fi abordate de științele naturii într-un mod evident. Prin urmare, presupunerea lui Bennett și Hacker că, în opinia mea, toate problemele filosofice pot fi în cele

din urmă transformate în probleme științifice cu soluții științifice dezvăluie o înțelegere greșită a argumentelor mele. Astfel de cazuri sunt, din punctul meu de vedere, excepționale.

Din nou, nu cred că este posibil să se facă o distincție cu adevărat clară și precisă, precum cea pe care pretind că au făcut-o, între întrebările empirice și cele conceptuale și, în consecință, nu fac o distincție clară între întrebările științifice și cele filozofice. Permiteți-mi să dau un exemplu pentru a explica cum pot ajuta descoperirile științifice munca mea filosofică. Când ridic brațul, intenția mea conștientă în acțiune **provoacă o mișcare fizică a corpului meu. Dar mișcarea are și un alt nivel de descriere, deoarece este cauzată de o secvență de descărcări neuronale și de secreția de acetilcolină la nivelul terminalelor axonale ale neuronilor motori. Pornind de la aceste fapte, pot face o analiză filosofică pentru a arăta că unul și același eveniment trebuie să fie atât un eveniment conștient calitativ, cât și subiectiv și să posede și multe proprietăți chimice și electrice. Dar aici se termină analiza filosofică . Acum trebuie să înțeleg mecanismul exact care face ca acest lucru să funcționeze, mai exact cum creierul provoacă și desfășoară intenția în acțiune, astfel încât intenția, cu combinația sa de structuri fenomenologice și electrochimice, să poată mișca obiecte fizice. Pentru aceasta, voi avea nevoie de rezultatele cercetării neurobiologice.

Intenția-în-acțiune *este* un concept introdus de Searle în analiza sa asupra acțiunii din **Intenționalitate (1983). Conform lui Searle, fiecare acțiune este compusă dintr-o intenție, care constituie cauza mentală, și o mișcare. În acțiunile premeditate, există o intenție prealabilă care provoacă actul. Intenția-în-acțiune este componenta intențională a unei acțiuni nepremeditate care, spre deosebire de intenția prealabilă, provoacă actul prin stabilirea din mers, ca să spunem așa, a condițiilor de satisfacție pe care o mișcare adecvată trebuie să le îndeplinească. Vezi Searle, J., *Intenționalitate: Un eseu în filosofia minții*, Cambridge, Cambridge University Press, 1983 (traducere în spaniolă: * Intencionalidad: Un ensayo en la filosofia *minții **, Madrid, Tecnos, 1992) .

Bennett și Hacker au scris o carte importantă și, în multe privințe, utilă. Au investit multă muncă în ea. Nu intenționez să pătez meritele lucrării lor cu obiecțiile mele. Cu toate acestea, cred că viziunea pe care o prezintă despre neurobiologie și minte este profund defectuoasă și potențial dăunătoare. Abordarea lor ar interzice multe dintre întrebările cruciale pe care trebuie să ni le punem în filosofie și neuroștiințe. De exemplu, ce sunt NCN-urile (corelatele neuronale ale conștiinței) și cum anume provoacă ele conștiința? Cum poate intenția mea în acțiune să-mi miște corpul? Dacă propunerile lui Bennett și Hacker ar fi acceptate, un număr vast de întrebări cheie din cercetarea neurobiologică ar fi respinse ca fiind lipsite de sens sau incoerente. De exemplu, întrebarea centrală din studiul neuroștiințific al vederii, „Cum provoacă experiențe vizuale conștiente procesele neurobiologice care încep cu atacul fotonilor asupra celulelor fotoreceptoare și continuă prin cortexul vizual în lobii prefrontali?", nu ar putea fi investigată de nimeni care a acceptat concepția autorilor noștri. Acesta este unul dintre acele cazuri, precum cel al inteligenței artificiale puternice, ††în care o teorie filosofică eronată poate avea consecințe științifice potențial dezastruoase și de aceea consider important să abordez tezele lui Bennett și Hacker .

††IA puternică (Inteligență Artificială) este o teză filosofică conform căreia ceea ce numim minți sunt de fapt programe digitale de calculator. Termenul de IA „puternică" este folosit pentru a o distinge de IA „slabă" sau „prudentă", care susține că computerele sunt instrumente foarte utile în studiul minții, dar nu că un computer are o minte în sens literal. Unele dintre corolăriile IA puternice sunt irelevanța neurofiziologiei pentru studiul minții și identificarea stărilor mentale și a stărilor computaționale. Searle își îndreaptă faimosul argument al Camerei Chineze împotriva acestei ultime teze. Vezi Searle, J., „Minți, Creiere și Programe" (1980), în *Behavioral and Brain Sciences ,* 3 (3), pp. 417-424 (traducere în spaniolă: „Minți, Creiere și Programe", capitolul 22 din Dennett, D. și Hofstader, D. (eds.), *The Mind's Eye: Fantasies and Reflections on the Self and the Soul,* Buenos Aires, Sudamericana, 1983). *(Notă a traducătorului)*

RĂSPUNSURI AU/S RESPUNZĂRI

Ipotezele conceptuale ale neuroștiinței cognitive

Răspuns la critici

MAXWELL BENNETT și PETER HACKER

CLARIFICARE CONCEPTUALĂ

Scopul nostru în *Fundamentele Filosofice ale Neuroştiinţei* [a fost] să contribuim la cercetarea în neuroștiințe în singurul mod în care filosofia poate ajuta știința: nu prin a oferi oamenilor de știință teorii empirice care să le înlocuiască pe ale lor, ci prin clarificarea cadrelor conceptuale pe care le utilizează. Unul dintre noi și-a dedicat viața construirii de teorii empirice despre funcțiile neuronale . Dar această activitate, care abordează fundamentele neuroștiinței , nu oferă niciunul dintre *fundamentele sale conceptuale.* Clarificările sistematice pe care le-am prezentat cu privire la senzație , percepție, cunoaștere, memorie, gândire, imaginație, emoție, conștiință și conștiință de sine nu sunt teorii. Scopul [lor] este de a clarifica conceptele psihologice pe care neurologii cognitivi le folosesc în teoriile *lor empirice.* Clarificările conceptuale pe care le-am prezentat demonstrează existența - numeroaselor inconsecvențe în teoretizarea actuală a neuroștiinței . Ele arată de ce se fac aceste greșeli și cum să le evităm.

Neuroștiința cognitivă este o investigație experimentală care își propune să descopere adevăruri empirice privind bazele neuronale

a facultăților umane și a proceselor neuronale care însoțesc exercitarea acestor facultăți. O precondiție a adevărului este sensul. Dacă o formă de cuvânt este lipsită de sens, nu poate exprima un adevăr, iar dacă nu exprimă un adevăr, nu poate explica nimic. Scopul cercetării filosofice asupra fundamentelor conceptuale ale neuroștiinței este de a dezvălui și clarifica adevărurile conceptuale - adevăruri pe care descrierile plauzibile ale descoperirilor și teoriilor neuroștiințifice le presupun și fără de care acestea sunt lipsite de sens. Dacă această cercetare este efectuată corect, va clarifica experimentele neuroștiințifice și descrierea lor, precum și inferențele care pot fi trase din ele. În Fundamentele *Filosofice ale Neuroștiinței,* delimităm rețeaua conceptuală formată din familii de concepte psihologice. Aceste concepte sunt presupuse de cercetarea neuroștiinței cognitive asupra bazei neuronale a puterilor cognitive, cogitative, afective și volitive umane. Dacă relațiile logice de implicație, excludere, compatibilitate și presupoziție care caracterizează utilizarea acestor concepte nu sunt respectate, este foarte probabil să se facă inferențe invalide, cele valide să fie trecute cu vederea, iar combinațiile de cuvinte fără sens să fie tratate ca și cum ar avea sens.

Unii filozofi, în special din Statele Unite, au fost puternic influențați de filosofia logicii și limbajului a lui Quine, conform căreia nu există o diferență semnificativă între adevărurile empirice și cele conceptuale .^[4] Astfel, din punct de vedere teoretic, un quinean ar susține că nu există o diferență esențială între, de exemplu, propoziția „memoria este cunoaștere reținută” și propoziția „memoria depinde de funcționarea normală a hipocampului și neocortexului”. Dar nu este cazul: prima propoziție exprimă un adevăr conceptual; a doua, o descoperire științifică. Potrivit lui Quine, afirmațiile unei teorii sunt confruntate cu experiența în întregime și sunt confirmate holistic. Însă este o greșeală să presupunem că propoziția „vulpile sunt femele” este confirmată de succesul teoriei zoologice sau că „burlacii nu sunt căsătoriți” de sociologia obiceiurilor conjugale. De asemenea, afirmația că roșul este mai închis decât rozul sau că roșul este mai asemănător cu portocaliul decât cu galbenul nu este verificată prin confirmarea din teoria culorilor , ci mai degrabă este o presupoziție a acestei teorii. Ar fi o greșeală

să presupunem că teoremele calculului diferențial au fost confirmate holistic de succesul predictiv al mecanicii newtoniene și slăbite de eșecul și respingerea acestei teorii. În realitate, ele au fost confirmate de *demonstrații matematice* . - Descoperirile sau teoriile empirice nu pot confirma sau slăbi propozițiile non-empirice, fie că sunt propoziții de logică, matematică sau adevăruri conceptuale manifeste.5 Adevărurile conceptuale definesc spațiul logic în care sunt situate faptele. Ele determină ce are sens. Prin urmare, faptele nu le pot nici confirma , nici intra în conflict cu ele.6

O propoziție conceptuală atribuie proprietăți sau relații interne ; o propoziție empirică atribuie proprietăți sau relații externe. Un adevăr conceptual este parțial constitutiv pentru semnificațiile expresiilor sale constitutive; o propoziție empirică este o descriere a modului în care sunt lucrurile. Un adevăr conceptual este o afirmație implicită a unei norme descriptive deghizată într-o afirmație factuală. Tocmai pentru că aceste tipuri de afirmații sunt parțial constitutive pentru semnificațiile expresiilor lor constitutive , incapacitatea de a recunoaște un adevăr conceptual (de exemplu, că roșul este mai închis decât rozul) este un criteriu indicativ al lipsei de înțelegere a uneia sau alteia dintre expresiile sale constitutive.

Este evident că această concepție normativă a adevărului conceptual nu se referă *în mod specific* la așa-numitele propoziții analitice în niciuna dintre diversele, familiare și *diferitele* concepții ale analiticității, cum ar fi cele ale lui Kant, Bolzano, Frege și Carnap. De fapt , diversele distincții dintre analitic și sintetic sunt pur și simplu ocolite. În schimb, distingem între afirmația unei măsuri și afirmația unei măsurători. A presupune că distincția pe care o facem între conceptual și empiric este o distincție *epistemică* ar fi un semn de neînțelegere totală.7 Această distincție *nu este* dedusă din *modul în care cunoaștem* adevărurile respective, ci din *rolul* propoziției în cauză: dacă este normativă (și constitutivă) sau descriptivă. Trebuie subliniat faptul că faptul că este una sau alta este o caracteristică a utilizării unui enunț, nu (sau nu neapărat) a unui tip de enunț. Aceeași enunț folosită într-un context pentru a exprima un adevăr conceptual poate fi folosită într-un alt context ca o afirmație de fapt - așa cum este bine demonstrat în mecanica newtoniană. În multe contexte, *fără investigații suplimentare,* s-ar

putea să nu fie clar ce rol ar trebui să joace o sintagmă în uz. Într-adevăr, în știință , este tipic ca dovezile inductive și constitutive (criteriile logice) să fluctueze. Dar este destul de clar că a caracteriza o sintagmă ca expresie a unui adevăr conceptual înseamnă a evidenția *funcția sa distinctivă* ca afirmație a unei măsuri, mai degrabă decât a unei măsurători. Prin urmare, distincția, spre deosebire de distincția *a priori posteriori ,* nu este epistemologică, ci logică sau logico-gramaticală.

Două PARADIGME: ARISTOTEL ȘI DESCARTES

Reflecția filosofică asupra naturii umane, asupra trupului și sufletului, datează din zorii filosofiei. Platon și Aristotel au stabilit polaritățile între care fluctuează această reflecție. Conform lui Platon și tradiției platonico-creștine a Sfântului Augustin, ființa umană nu este o substanță unificată, ci o combinație a două - substanțe distincte: un corp muritor și un suflet nemuritor. Conform lui Aristotel , ființa umană *este* o substanță unificată, în care sufletul *(psihicul)* este forma corpului. A descrie această formă înseamnă a descrie capacitățile caracteristice ale ființelor umane, în special capacitățile distinctive ale intelectului și voinței care caracterizează *psihicul rațional* . Dezbaterea modernă pe această temă începe cu moștenitorul tradiției platonico-augustiniene, și anume, concepția cartesiană a ființelor umane ca două lucruri într-unul: o minte și un corp. Interacțiunea lor cauzală reciprocă a fost invocată pentru a explica experiența și comportamentul uman.

Cele mai importante figuri ale primelor două generații de neurocercetători din secolul al XX-lea - de exemplu, Sherrington, Eccles și Penfield - au fost dualiști cartezieni declarați. A treia generație a păstrat structura cartesiană de bază, dar a transformat-o în dualism corp-creier: dualismul substanțial a fost abandonat, dar dualismul structural a fost menținut. Neurocercetătorii contemporani atribuie creierului același set de predicate mentale pe care Descartes le-a atribuit minții și concep în mare măsură relația dintre gândire și acțiune și dintre experiență și

obiectele sale în același mod în care a făcut-o Descartes, cu excepția , în esență, a simplei înlocuiri a minții cu creierul . Tema centrală a cărții noastre a fost să demonstrăm incoerența dualismului creier/corp și să dezvăluim caracterul său cripto- cartezian eronat. Partea constructivă a scopului nostru a fost să demonstrăm că, pentru a face dreptate structurii schemei noastre conceptuale și pentru a oferi descrieri coerente ale marilor descoperiri ale neuroștiinței cognitive post-Sherrington, este necesară o explicație aristotelică, cu accent pe capacitățile active și pasive de ordinul întâi și doi și pe manifestările lor comportamentale, completată de perspective wittgensteiniene care se adaugă celor ale lui Aristotel.8

PRINCIPIUL LUI ARISTOTEL ȘI EROAREA MEREOLOGICĂ

În *lucrarea „Fundamentele filosofice ale neuroștiinței", semnalăm o* eroare persistentă pe care o numim „eroarea mereologică a neuroștiinței ".9 Corectarea acestei erori este un laitmotiv (dar *numai* un laitmotiv) al cărții noastre. Spunem că eroarea este „mereologică" deoarece implică atribuirea părților de atribute care pot fi atribuite în mod inteligibil doar întregurilor din care fac parte. Aristotel a semnalat o formă a acestei erori în jurul anului 350 î.Hr., când a observat că „a afirma [...] că sufletul este cel *care* este mânios ar fi ca și cum ai afirma că sufletul este cel care țese sau construiește. Ar fi mai bine, de fapt , să nu spunem că sufletul este cel care se compătimește, învață sau raționează, ci omul în virtutea sufletului" (DA 408 ^{b} 12-15) ‡‡- a face ceva în virtutea sufletului cuiva este același lucru cu a-l face în virtutea abilităților cuiva. Este o greșeală să atribui sufletului unui animal atribute care pot fi atribuite doar literal animalului ca întreg. Putem numi acest principiu „principiul lui Aristotel".

Interesul nostru principal a fost în echivalentul neuroștiințific al acestui principiu, și anume, eroarea de a atribui creierului - *unei* părți a unui animal -

‡‡Aristotel, *Despre suflet,* Madrid, Gredos, 1978, p. 155, trad. de Tomás Calvo Martínez. *(Nota traducătorului)*

atribute care pot fi atribuite animalului doar ca întreg. Nu am fost primii care au subliniat acest lucru; Anthony Kenny făcuse acest lucru în strălucitul său articol din 1971, „Eroarea homunculului”.^[10] Această eroare este mai degrabă *mereologică* decât strămoșul său aristotelic, deoarece creierul este literalmente o parte a animalului simțitor, în timp ce, contrar afirmațiilor lui Platon și Descartes, sufletul sau mintea nu sunt. Astăzi , într-un spirit aristotelic, observăm că a spune că creierul este iritat este ca și cum ai spune că creierul țese sau construiește. Într-adevăr, este cu siguranță mai bine să nu spunem că creierul se compătimește, învață sau raționează, ci mai degrabă că o persoană face toate aceste lucruri. [11] Prin urmare, negăm faptul că are sens să spunem că creierul este conștient, are senzații, percepe, gândește, știe sau dorește orice, întrucât acestea sunt atribute ale animalelor, nu ale creierului lor.

Am fost puțin surprinși că profesorul Dennett a crezut că ceea ce *aveam* în minte era distincția stabilită în lucrarea sa *„Content and Conscious” (Conținut și conștientizare).* din 1969*, între nivelurile personal și subpersonal ale explicațiilor. Acolo el a afirmat, corect, că a simți durerea nu este o proprietate a creierului. Dar motivul pe care l-a dat a fost că durerile sunt „fenomene mentale” care sunt „nemecanice”, în timp ce procesele cerebrale sunt „în esență mecanice”. [12] Contrastul pe *care l-am* stabilit între proprietățile întregului și cele ale părților nu este cel dintre nemecanic și mecanic . Ceea ce indică ora este ceasul ca întreg, nu axele sau roțile sale dințate - deși procesul de a spune ora este în întregime - mecanic. Ceea ce zboară este avionul, nu motoarele sale - deși procesul de zbor este în întregime mecanic. Mai mult, verbele senzaționale, cum ar fi „a răni”, „a mănânca” sau „a gâdila”, *se* aplică părților unui animal al cărui picior îl poate răni, al cărui cap îl poate mănânca sau ale cărui flancuri pot fi gâdilate (PFN, p. 73). Aceste atribute sunt, în cuvintele profesorului Dennett, „nemecanice”;

Traducere în spaniolă: *Conținut și conștiință,* Barcelona, Gedisa, 1996. *(Nota traducătorului)* însă, *acestea sunt* atribuibile unor părți ale animalului. Așadar, observația noastră mereologică este foarte diferită de distincția pe care profesorul Dennett o stabilește între nivelurile de explicație personală și subpersonală și, aplicată animalelor, foarte diferită de distincția sa dintre ceea ce este „mecanic” și

ceea ce nu este. [13]

ESTE EROAREA MEREOLOGICĂ ÎNTR-ADEVĂR MEREOLOGICĂ?

Profesorul Searle a obiectat că ceea ce am caracterizat drept paradigmă a unei erori mereologice – adică atribuirea de atribute psihologice creierului – nu este așa, deoarece creierul nu este o parte a persoanei, ci mai degrabă o parte a corpului unei persoane (p. 107). Aceasta, în opinia noastră, este pur și simplu o încurcătură. Cuvintele lui Wittgenstein pe care le-am citat au fost: „Numai despre *ființele umane* și despre lucrurile care seamănă cu ființele umane (se comportă ca ființele umane) se poate spune: au senzații; văd, sunt orbi; aud, sunt surzi; sunt conștienți sau inconștienți" [14] (sublinierea noastră). Creierul *este* o parte a ființei umane.

Profesorul Searle sugerează că, dacă atribuirea de atribute psihologice creierului ar fi într-adevăr o eroare mereologică, o astfel de eroare ar dispărea dacă acele atribute ar fi atribuite la ceea ce el numește „restul sistemului" căruia îi aparține creierul. El consideră că „restul sistemului" este corpul uman. El observă că nu atribuim predicate psihologice corpului pe care îl *avem* . Cu excepția notabilă a verbelor care indică senzația (de exemplu, „Mă doare tot corpul"), ultimul punct este corect. Nu spunem: „Corpul meu percepe , gândește sau știe". Cu toate acestea, „sistemul" căruia se poate spune că îi aparține creierul uman este *ființa umană.* Creierul uman este o parte a ființei umane, la fel cum creierul canin este o parte a câinelui. Creierul meu - creierul pe care îl am - este o parte din mine - din ființa umană vie care sunt - la fel cum sunt picioarele și brațele mele. Dar este, de asemenea, adevărat că se poate spune că și creierul meu este o parte a corpului meu.

Cum poate fi explicat acest lucru? Când vorbim despre mintea noastră, o facem în mare parte într-un *sens non-agentiv* și metaforic, referindu-ne la capacitățile noastre raționale de intelect și voință și la exercitarea acestora. Când vorbim despre corp, vorbim despre proprietățile noastre corporale. A vorbi despre *corpul meu* înseamnă a vorbi despre trăsăturile corporale *ale ființei umane care*

sunt - trăsături care țin de aspect (un corp atractiv sau neîndemânatic), de suprafețele ființei umane (corpul său era acoperit de mușcături de țânțari, lacerat peste tot, vopsit în albastru), de aspectele sănătății și stării fizice (un corp sănătos sau bolnav) și, în mod surprinzător , de senzație (tot corpul meu poate durea, așa cum mă doare piciorul sau mă mănâncă spatele). Dar a cunoaște, a percepe, a gândi, a imagina și [așa] mai departe nu sunt caracteristici corporale ale ființelor umane și nu sunt atribuibile corpului pe care îl are ființa umană, așa cum nu sunt atribuibile creierului pe care îl are ființa umană. Ființele umane nu sunt *corpurile lor.* Cu toate acestea, *ele sunt* corpuri, în sensul complet diferit de a fi un anumit tip de ființă spatiotemporală persistentă, sensibilă - *Homo sapiens* - iar creierul este o parte a ființei umane vii, la fel ca și membrele.16 Cu toate acestea, nu este o parte conștientă, gânditoare și perceptivă , așa cum nicio altă parte a ființei umane nu este. Acestea sunt atribute ale ființei umane ca întreg.

Totuși, profesorul Searle a subliniat o caracteristică interesantă a limbajului nostru metaforic referitor la corp. Ființele umane sunt persoane - adică animale inteligente capabile să folosească limbajul - sunt conștiente de sine, posedă cunoștințe despre bine și rău, sunt responsabile pentru acțiunile lor și au drepturi și obligații . A fi persoană, simplu spus, înseamnă a poseda acele capacități care conferă statutul de agent moral. Probabil că nu am spune că creierul face parte din persoană, ci mai degrabă că face parte din corpul persoanei, în timp ce nu am ezita să spunem că creierul lui Pepe este o parte din Pepe, o parte din *această* ființă umană, în același mod, exact, în care picioarele și brațele lui Pepe sunt părți ale lui Pepe. De ce? Poate pentru că „persoană" este, așa cum a subliniat Locke , „un termen criminalistic", dar nu o substanță de nume. Prin urmare, dacă folosim termenul „persoană" în contexte ca acesta, indicăm că ceea ce ne interesează în primul rând sunt ființele umane în măsura în care posedă acele caracteristici care le fac persoane, relativ independente de caracteristicile lor fizice . Următoarea analogie ar putea fi de ajutor: Londra face parte din Regatul Unit; Regatul Unit aparține Uniunii Europene, dar Londra nu. Acest lucru nu împiedică Londra să facă parte din Regatul Unit . În mod similar, faptul că Pepe este o persoană nu

împiedică creierul său să facă parte din el.

BAZA PRINCIPIULUI

De ce ar trebui să acceptăm principiul lui Aristotel și echivalentul său neuroștiințific? De ce ar trebui să-i descurajăm pe neurocercetătorii care insistă să atribuie creierului *conștiința, cunoașterea, percepția etc. ?*

Conștiință . Animalele sunt conștiente sau inconștiente și pot deveni conștiente de ceva ce le captează atenția. Elevul, nu creierul său, este cel care se trezește și devine conștient de ceea ce spune profesorul și este...

Profesorul, nu creierul său, este conștient de plictiseala elevilor săi văzându-i căscând discret. Creierul nu este organul conștiinței. Vedem cu ochii și auzim cu urechile, dar nu suntem conștienți cu creierul, așa cum nu mergem cu el.

Un animal poate fi conștient fără a demonstra acest lucru. Acesta este *singurul sens* în care se poate spune, împreună cu profesorul Searle, că „însăși existența conștiinței nu are nicio legătură cu comportamentul ” (p. 104). Însă *conceptul* de conștiință este strâns legat de dovezile comportamentale pentru atribuirea conștiinței unui animal . Un animal nu trebuie să prezinte un anumit comportament pentru *a fi* conștient . Dar numai despre un animal căruia *i se poate atribui în mod inteligibil acel comportament* se poate afirma că este conștient și că această afirmație poate fi *adevărată sau falsă.* Nu are sens să atribuim conștiință sau raționament unui scaun sau unei stridii, deoarece nu există niciun scaun sau nicio stridie care să adoarmă și apoi să se trezească sau care să-și piardă conștiința și să o recapete; și nu există niciun scaun sau nicio stridie care să se comporte rezonabil sau nerezonabil. [17] „Întrebarea ontologică” (așa cum o numește profesorul Searle) - întrebarea adevărului (așa cum preferăm să o numim) - presupune determinarea prealabilă a întrebării sensului. Acordul asupra dovezilor comportamentale pentru atribuirea conștiinței, adică asupra a ceea ce *contează* drept manifestare a

conștiinței, este o condiție prealabilă pentru cercetarea științifică a cerințelor neuronale pentru conștiință. Altfel, nici măcar nu ar fi posibil să se identifice ce trebuie investigat. A distinge chestiunea sensului de chestiunea adevărului nu înseamnă a confunda „regulile de utilizare a cuvintelor cu ontologia", așa cum subliniază profesorul Searle (p. 105); dimpotrivă, înseamnă a le diferenția. [18]

Profesorul Searle insistă că conștiința este o proprietate a creierului. Sherrington, Eccles și Penfield, fiind cartezieni, au crezut în mod eronat că este o proprietate a minții. Ce experiment neuroștiințific recent poate oferi profesorul Searle pentru a demonstra că este *într-adevăr* o proprietate a creierului? La urma urmei, singurul lucru pe care neurocercetătorii *l-au putut* descoperi este că anumite stări neuronale sunt corelate inductiv cu conștiința unui *animal* și sunt condiții cauzale ale acesteia. Dar *această* descoperire nu poate dovedi că *creierul* este conștient. Este, așadar, afirmația profesorului Searle o reflecție conceptuală? Nu - pentru că nu așa este folosit conceptul de *conștiință* . Ființele umane (și alte animale), nu creierul lor (sau mintea lor), care adorm și apoi se trezesc, pot fi doborâte inconștiente și apoi își pot recăpăta conștiința. Prin urmare, este o - recomandare lingvistică, în mod specific, ca atunci când un creier uman se află într-o stare corelată inductiv cu faptul că ființa umană este conștientă, ar trebui să descriem și creierul său ca fiind conștient? Este o convenție pe care am putea- o adopta. Am putea introduce această utilizare derivată a cuvântului „a fi conștient". Ar depinde în mod necesar de utilizarea principală care se aplică ființei umane în ansamblu. Cu toate acestea, este dificil să se găsească vreun motiv pentru a o face recomandabilă. Cu siguranță nu contribuie cu nimic la claritatea descrierii și nu adaugă nimic altceva decât o formă goală explicației neuroștiințifice existente.

Cunoaștere . Cunoașterea cuprinde capacități de diferite tipuri. Identitatea unei capacități este determinată de ceea ce reprezintă. Cel mai simplu motiv pentru atribuirea unei capacități unui animal este faptul că acesta desfășoară activități corporale care manifestă acea capacitate. Cu cât capacitatea este mai complexă, cu atât motivele pentru atribuirea ei sunt mai diverse și mai difuze. Dacă un animal știe ceva, acesta poate acționa și reacționa la mediul său în moduri în care ar fi

incapabil să le facă dacă nu l-ar cunoaște; dacă manifestă aceste moduri în comportamentul său, își demonstrează cunoștințele. S-ar putea spune că creierul este *vehiculul* acestor capacități, dar aceasta înseamnă că, în absența structurilor neuronale adecvate, animalul nu ar fi capabil să facă ceea ce face. Structurile neuronale ale creierului sunt diferite de capacitățile *pe care le are animalul,* iar funcționarea acestor structuri este diferită de exercitarea acestor capacități *de către animal.* Pe scurt, cel care știe este și cel care acționează, iar cunoștințele sale se manifestă în acțiunile sale.

Îl criticăm pe J.Z. Young pentru că susține, așa cum fac mulți neurocercetători , că creierul conține cunoștințe și informații „în același mod în care cunoștințele și informațiile sunt înregistrate în cărți sau computere". ¹⁹ Profesorul Dennett afirmă că nu am făcut nimic pentru a stabili că nu există niciun concept de cunoaștere sau informație despre care să nu se poată spune că este codificat atât în cărți, cât și în creier (p. 91). De fapt, am examinat problema (PFN, pp. 152 și următoarele). Dar o vom explica din nou.

Un cod este un sistem de convenții pentru codificarea și transmiterea - informațiilor care depinde de limbă. Un cod nu este o limbă; nu are gramatică sau lexic (de exemplu, codul Morse). Cunoașterea nu este *codificată* în cărți decât dacă este scrisă în cod. Un mesaj poate fi codificat doar dacă există un cod în care să se facă acest lucru, iar un cod există doar dacă codificatorii și decodificatorii cărora li se adresează sunt de acord asupra convențiilor de codificare. În acest sens, un cod neural nu există și nu ar putea exista. Creierul nu conține informații așa cum o face o carte. Creierul nu posedă informații așa cum o face o ființă umană. Faptul că informațiile pot fi derivate din caracteristicile creierului (la fel cum informațiile dendrocronologice pot fi derivate din trunchiul unui copac) nu dovedește că informațiile sunt codificate în creier (așa cum nu sunt codificate în trunchiul copacului).

Prin urmare, în sensul obișnuit al termenului „cunoaștere", creierul nu poate înregistra, conține sau deține cunoștințe. În acest moment , profesorul Dennett schimbă tactica și recomandă să analizăm literatura de științe cognitive privind

extensiile termenului „cunoaștere” care ar putea permite atribuirea cunoașterii, într-un sens extins, creierului. El recomandă să analizăm încercarea lui Chomsky de a explica un concept extins de cunoaștere, și anume cel de „cogniție”, conform căruia ființele umane, chiar și nou-născuții, [cunosc] *principiile* gramaticii universale. Potrivit lui Chomsky, cineva care *cunoaște* nu poate spune altuia ceea ce *cunoaște* , nu poate arăta obiectul *cunoașterii sale,* nu recunoaște ceea ce *cunoaște* atunci când i se descrie, nu uită niciodată ceea ce *cunoaște* (dar nici nu-și amintește vreodată), nu l-a învățat niciodată și nu l-ar putea preda. În afară de aceasta, *a cunoaște* este același lucru cu a cunoaște. Este *acesta postulat* ca model al unei extensii inteligibile a unui termen?

Percepția . Facultățile perceptive sunt abilitățile de a dobândi cunoștințe prin utilizarea propriilor sisteme senzoriale . Animalul își folosește ochii pentru a privi, a contempla, a zări și a observa lucruri. În acest fel, este capabil să distingă lucruri care au culori, forme și mișcări distinctive. Își demonstrează acuitatea vizuală prin acțiunile sale ca răspuns la ceea ce vede. Nu ar poseda aceste abilități perceptive și nici nu ar fi capabil să le exercite dacă nu ar exista o funcționare corectă a părților corespunzătoare ale creierului său. Cu toate acestea, nu cortexul cerebral vede, ci animalul. Nu creierul este cel care se apropie de lucruri pentru a le vedea mai bine, cel care privește prin tufișuri și pe sub garduri vii. Nu creierul este cel care sare să fugă de un prădător pe care l-a detectat și nici cel care atacă prada pe care a observat-o; animalul înzestrat cu percepție face toate acestea. Pe scurt, cel care percepe este și cel care acționează.

În *lucrarea „Conștiința explicată”,* profesorul Dennett atribuie creierului atribute psihologice. El afirmă că acesta este conștient, adună informații, face presupuneri simple, se bazează pe informațiile disponibile și ajunge la [concluzii.21]

Aceasta înseamnă căderea în aceeași eroare de rațiune împotriva căreia au avertizat atât Aristotel, cât și Wittgenstein: eroarea mereologică, așa cum am numit-o. În articolul său pentru APA, profesorul Dennett admite că ar fi o eroare să atribuim predicate psihologice *complete* unor părți ale creierului (p. 87). Cu toate acestea, el susține că este teoretic util și consistent cu acceptarea naturii

eronate a atribuirii predicatelor întregului părților sale să extindem vocabularul psihologic , *atenuat în mod corespunzător,* al oamenilor și al altor animale a) computerelor și b) părților creierului. Într-adevăr, el susține deschis că nu există o diferență semnificativă între aceste două extensii. Există însă o diferență: atribuirea de proprietăți (atenuate sau nu) computerelor este o eroare, dar nu constituie o eroare mereologică. A considera creierul ca un computer și a-i atribui astfel de proprietăți psihologice, fie ca întreg, fie părților sale, este dublu eronat. Să explicăm.

Este adevărat că, în vorbirea informală, spunem că calculatoarele își amintesc, că își caută memoria, că calculează și, uneori, când le ia mult timp, spunem în glumă că gândesc. Dar aceasta este doar o figură de stil, nu o aplicare literală a termenilor „a-și aminti", „a calcula" și „a gândi". Calculatoarele sunt dispozitive concepute pentru a îndeplini anumite funcții în beneficiul nostru. Putem stoca informații pe un computer, la fel cum putem face într-un fișier. Dar fișierele nu își pot aminti nimic, la fel cum nici computerele nu pot. Folosim computerele pentru a obține rezultatul unui calcul - la fel cum foloseam o riglă de calcul sau un calculator mecanic . Aceste rezultate sunt obținute fără ca cineva sau ceva să calculeze la propriu - așa cum este evident în cazul unei rigle de calcul sau al unui calculator mecanic. Pentru a calcula la propriu, trebuie să stăpânești o mare varietate de concepte, să urmezi o multitudine de reguli pe care trebuie să le cunoști și să înțelegi o gamă întreagă de operațiuni. Calculatoarele nu fac acest lucru și nici nu pot.

Profesorul Dennett sugerează că „este un fapt empiric [...] că *anumite părți* ale creierului nostru desfășoară procese *izbitor de* similare în conjecturare, luarea de decizii, credința, tragerea concluziilor și așa mai departe. Și sunt *suficient de similare* cu aceste comportamente la nivel personal pentru a justifica extinderea limbajului obișnuit pentru a le cuprinde" (p. 86). El este de acord că ar fi o greșeală să „atribuim credințe", decizii, dorințe sau chiar suferință *creierului* . Este mai bine să spunem că „așa cum un copil mic poate crede *ceva de genul* că tatăl său este doctor [...] o anumită parte a creierului unei persoane poate crede *ceva de genul* că

există o ușă deschisă la câțiva metri înainte” (p. 87).

Aceasta face parte din ceea ce profesorul Dennett caracterizează drept „ - atitudinea intențională” - o metodologie de cercetare care se presupune că îi ajută pe neurocercetători să explice baza neuronală a capacităților umane. Dennett spune că adoptarea atitudinii intenționale a permis „o muncă științifică excelentă [...] generarea de ipoteze testabile, dezvoltarea de teorii, analizarea fenomenelor extrem de complexe în părțile lor cele mai inteligibile ” (p. 87). El pare să accepte ideea că unele părți ale creierului „cred într-un fel”, altele *decid într-un fel, iar* altele *monitorizează într-un fel* aceste activități. Se presupune, probabil , că toate aceste *lucruri explică* ceea ce vor neurocercetătorii să explice. Dar dacă ceea ce se explică este un fel de credințe, pseudo-așteptări, proto-dorințe și semi-decizii (așa cum sugerează Dennett [p. 88]), în cel mai bun caz, acestea reușesc doar să aibă un oarecare *sens* și se poate presupune că sunt doar *într-un oarecare mod adevărate.* Și modul în care se pot trage inferențe valide din aceste premise este mai mult decât obscur. Modul exact în care astfel de premise ar trebui *să explice* fenomenele este la fel de obscur. Logica acestor pretinse explicații este complet confuză. Oferă așa-numitele credințe, pseudo-credințe, proto-credințe sau semi-credințe despre ceva unei părți a creierului un motiv pentru a acționa? Sau doar așa-numitul motiv - pentru așa-numita acțiune? Când Dennett este întrebat dacă părțile creierului sunt, așa cum susține el, „ sisteme intenționale autentice”, răspunsul său este: „Nu întrebați” (p. 88). [22]

Neurocercetătorii cognitivi pun întrebări *reale : cum* este implicat cortexul prefrontal în gândirea umană, *de ce* există căi de reintrare *și care* sunt exact rolurile respective ale hipocampului și neocortexului în memoria umană. A spune că hipocampul își amintește cumva pentru scurte momente, iar neocortexul, pe de altă parte, are ceva asemănător memoriei pe termen lung, nu este deloc o explicație. Nicio teorie confirmată empiric în neuroștiințe nu a apărut din explicațiile lui - Dennett , deoarece atribuirea de „a ceva de genul proprietăților psihologice” unor părți ale creierului nu *explică* nimic. Vom reveni la acest subiect când vom discuta explicația lui Sperry și Gazzaniga asupra comisurotomiei. Nu numai că nu oferă

nicio explicație , dar generează și mai multă inconsecvență. [23]

Suntem de acord cu profesorul Dennett că multe dintre convingerile unui copil sunt atenuate. Înțelegerea conceptului de doctor de către o fetiță poate fi defectuoasă , dar ea va spune corect: „Tatăl meu este doctor", iar la întrebarea: „Unde este doctorul?", ea va răspunde: „Acolo" (arătând spre cabinetul tatălui ei). Astfel, se poate spune că fata crede, într-un sens atenuat, că tatăl ei este doctor. În comportamentul ei verbal și deictic, fata îndeplinește *unele* dintre criteriile normale pentru a crede că tatăl ei este doctor (dar îndeplinește și unele dintre criteriile pentru a nu crede că el este). Dar nu există așa ceva ca o parte a creierului care afirmă lucruri (așa cum face fata), care răspunde la întrebări (așa cum face fata) sau care indică lucruri deictic (așa cum face fata). Astfel, în sensul că fata, în comportamentul ei verbal și deictic, poate exprima o *credință rudimentară,* o parte a creierului nu poate, la fel cum creierul în ansamblu nu poate exprima o credință completă. Sau poate profesorul Dennett să sugereze un experiment crucial care să demonstreze că, de exemplu, cortexul prefrontal al fetei crede că pisica este sub canapea?

Copilul poate manifesta, de asemenea, o credință rudimentară în comportamentul său nonverbal. Dacă vede pisica alergând să se ascundă sub canapea și ea, cu mersul ei încă nesigur, merge să o caute acolo, s-ar putea spune că ea crede că pisica este sub canapea. Dar creierul și părțile sale nu se pot *comporta* în niciun fel; nu se pot apropia de canapea cu un mers nesigur, nu se pot uita sub ea și nu pot părea nedumeriți dacă nu există nicio pisică sub canapea . Părțile creierului nu pot acționa voluntar sau întreprinde acțiuni cu un scop în minte. Spre deosebire de copil, părțile creierului nu pot îndeplini *niciunul* dintre criteriile pentru a crede ceva, nici măcar într-un sens rudimentar. Creierul (și părțile sale) pot doar „oarecum să creadă" în același sens în care sunt „oarecum ca oceanele" (deoarece există undele cerebrale) și „oarecum ca sistemele meteorologice" (deoarece există brainstorming-ul). Există la fel de multă asemănare între procesele cerebrale și credințele, deciziile sau conjecturile umane precum există între creier și mare. La urma urmei, atât creierul, cât și marea sunt

gri, au riduri la suprafață și curenți care le străbat.

LOCALIZAREA ATRIBUTELOR PSIHOLOGICE

Întrebarea dacă creierul este un posibil subiect al atributelor psihologice este distinctă de cea dacă este locul acelor atribute psihologice cărora li se poate atribui în mod inteligibil o locație corporală (PFN, pp. 122 și următoarele și 179 și următoarele). Motivele care ne-au determinat să negăm că creierul poate fi subiectul atributelor psihologice nu demonstrează că acesta nu este locul acelor atribute cărora are sens să le atribuim o locație corporală. Nici nu au intenționat să o facem. Din perspectiva noastră, *este posibil* să atribuim o locație senzațiilor precum durerea și mâncărimea. Locația durerii este acolo unde persoana care suferă indică faptul că se află, în membrul masat, în partea corpului despre care spune că doare - deoarece aceste forme de comportament dureros oferă criteriile pentru localizarea sa. În schimb, gândirea, credința, decizia și dorința, de exemplu, nu pot primi o *locație somatică.* Răspunsul la întrebările „De unde ți-a venit ideea asta?", „De unde ai dobândit această credință ciudată?", „De unde ai luat decizia de a te căsători?" nu este *niciodată* „În cortexul prefrontal, desigur". Criteriile - pentru a determina unde o ființă umană s-a gândit la ceva, a dobândit o credință, a luat o decizie, s-a înfuriat sau a fost uimită implică un comportament, cu siguranță, dar nu un comportament care indică o localizare somatică. Locația gândirii, amintirii, vederii, deciziei, enervării sau uimirii unei ființe umane [este] oriunde *se află ființa umană atunci când gândește* etc. Ce parte a creierului său este implicată în efectuarea acestor lucruri este o altă întrebare importantă pe care neurologii o înțeleg treptat mai bine. Dar ei nu descoperă unde are loc gândirea, amintirea sau decizia; ei descoperă ce părți ale cortexului sunt implicate cauzal în gândirea, amintirea sau decizia umană.

Desigur, a te gândi la ceva, a decide să faci ceva, a vedea ceva sunt, așa cum spune pe bună dreptate profesorul Searle (p. 110), evenimente reale; ele se

întâmplă de fapt undeva în lume, la un moment dat. Am venit cu acest argument în bibliotecă și am decis cum să-l formulez în biroul meu; l-am văzut pe Pepe când era pe stradă și am ascultat recitalul lui Inés în auditoriu. Profesorul Searle sugerează că întrebarea „Unde se întâmplă evenimentele mentale?" nu este mai abstrusă din punct de vedere filosofic decât întrebarea „Unde au loc procesele digestive?". Procesele digestive, susține el , au loc în stomac, iar conștiința are loc în creier. Aceasta este o greșeală. A fi conștient, spre deosebire de a fi inconștient, a fi conștient de ceva, spre deosebire de a nu-l observa sau de a nu-i acorda atenție, nu au loc *în* creier în niciun fel. Ele se întâmplă, desigur, *din cauza anumitor evenimente din creier,* fără de care ființele umane nu și-ar fi recăpătat conștiința și nici atenția lor nu ar fi fost atrasă asupra ei. Pentru a răspunde la întrebarea „Unde ați devenit conștient de sunetul ceasului?" Trebuie să specific unde mă aflam când un astfel de sunet mi-a atras atenția, la fel cum pentru a răspunde la întrebarea „Unde ți-ai recăpătat cunoștința?" trebuie să specific unde mă aflam când mi-am revenit.

Atât digestia, cât și gândirea sunt caracteristice animalelor, dar nu rezultă că nu există diferențe logice între ele. Se poate spune că stomacul digeră mâncarea, dar nu că creierul gândește. Stomacul este organul digestiv, dar creierul nu este un organ al gândirii ' la fel cum nu este un organ al locomoției. Dacă stomacul este deschis, se poate vedea cum are loc digestia în el. Dar dacă cineva vrea să vadă cum are loc gândirea, ar trebui să observe *Gânditorul* de Rodin (sau chirurgul în timpul operației , jucătorul de șah în timpul jocului sau dezbaterul în fierbințeala momentului), nu creierul. Tot ce poate arăta creierul unei persoane este ce se întâmplă acolo *în timp ce persoana gândește ;* tot ce pot arăta imaginile RMN funcțional este care parte a creierului acelei persoane metabolizează mai mult oxigen decât altele atunci când pacientul în cauză gândește. [26] (Atribuim lungimea, rezistența și prezența fisurilor grinzilor de fier. Dar nu rezultă din aceasta că *lungimea* și *rezistența* au același caracter logic; și putem spune unde este fisura, dar nu unde este rezistența.)

Astfel, senzațiile, cum ar fi durerea, *sunt* localizate în corpul nostru. Însă

profesorul Searle susține că toate se află *în creier.* El însuși recunoaște că acest lucru contravine intuiției - la urma urmei, ne plângem de dureri de stomac, gută la picioare sau artrită la genunchi. Cu toate acestea, afirmă el, creierul creează o imagine corporală, iar durerea despre care spunem că este în piciorul nostru și pe care o ameliorăm prin masarea acestuia este o conștiință a durerii așa cum este localizată în picior, care se găsește în imaginea corporală din creierul nostru. Interesant este că Descartes a adoptat o concepție foarte similară, menționând că „durerea din mână este simțită de suflet nu pentru că este prezentă în mână, ci pentru că este prezentă în creier". [27] Avantajul explicației sale , subliniază profesorul Searle, este că înseamnă că putem descrie fenomenul durerii fantomă fără a cădea în absurditatea de a sugera că durerea este în spațiul fizic, în pat sau sub cearșaf. Cădem în aceeași absurditate, argumentează el, atunci când spunem că durerile sunt în corp. Suntem de acord că este absurd, dar nu că ar trebui să cădem în asta.

Există multe utilizări locative ale cuvântului „în", unele spațiale, altele non-spațiale („în poveste", „în octombrie", „în comitet"). Printre utilizările spațiale, există multe tipuri diferite, în funcție de ce este în ce (PFN, pp. 123 și următoarele). Suntem de acord cu profesorul Searle că, dacă există o monedă în buzunarul jachetei mele și jacheta mea este în dulap, atunci există o monedă în dulap. Dar nu toate utilizările locative ale cuvântului „în" sunt tranzitive în acest sens. Dacă există o gaură în jacheta mea și jacheta este în dulap, nu rezultă că există o gaură în dulap. În cazul jachetei și al monedei, avem de-a face cu relații spațiale între două obiecte independente, dar nu este cazul jachetei și al găurii. În mod similar, dacă există o cută în cămașa mea și cămașa mea este în valiză, nu rezultă că există o cută în valiză. Moneda poate fi scoasă din buzunarul jachetei , iar cămașa poate fi scoasă din valiză, dar gaura nu poate fi scoasă din buzunar, trebuie reparată, la fel cum șifonul trebuie călcat, nu scos.

Utilizarea cuvântului „înăuntru" în ceea ce privește localizarea senzațiilor nu este ca în cazul monedei, ci mai degrabă ca în cazul găurii (deși, în orice caz, diferită). Durerea nu este o substanță. Dacă simt durere în picior, nu mă plasez în

nicio *relație* cu o durere - dimpotrivă, piciorul meu doare *acolo* și pot indica locul care doare, pe care îl numim „locația durerii". În cazul durerii de membru fantomă, persoana care o experimentează simte ca și cum membrul amputat ar fi încă la locul lui și raportează durere în membrul iluzoriu . Simte ca și cum o doare piciorul, în ciuda faptului că nu are picior. Suntem de acord cu profesorul Searle că nu patul doare și că durerea pe care o simte amputatul nu este sub cearșaf. Faptul că simte durere acolo unde ar fi piciorul său și că piciorul său ar fi sub cearșaf nu implică faptul că există durere sub cearșaf, la fel cum faptul că are durere în - piciorul neamputat și că acel picior este în cizmă nu implică faptul că are durere în cizmă. De fapt, suntem de acord cu profesorul Searle în ceea ce privește fenomenul și nu suntem de acord doar cu descrierea sa . Faptul că durerea persoanei amputate este reală, dar că locul unde o simte este iluzoriu (piciorul nu-l doare, pentru că nu are picior), nu dovedește că atunci când o persoană care *nu a* suferit o amputare simte durere în picior, locul unde o simte este, de asemenea, iluzoriu. Într-adevăr, piciorul este cel care doare. Nu credem că există imagini ale corpului în creier și ne întrebăm ce dovezi există pentru existența lor - la urma urmei, dacă deschizi un creier uman, nu vei găsi nicio imagine a corpului în el. La ce pare să se refere profesorul Searle este că metodele fiziologice, începând cu cea a lui Sherrington , au fost folosite pentru a stabili că neuronii din cortexul somatosenzorial pot fi excitați într-o relație unu-la-unu cu puncte stimulate de pe suprafața corpului și cu distribuția spațială a mușchilor membrelor și trunchiului. Însă este complet confuz ce vrea să spună profesorul Searle când vorbește despre „a avea o durere fantomă fenomenologică de picior în imaginea corporală care există în creier" (p. 118 și următoarele). Se *pot* avea dureri de *cap,* cunoscute în mod obișnuit sub numele de durere de cap. Dar nu se pot avea dureri de spate sau de stomac în creier sau orice alt tip de durere. Și aceasta nu este o coincidență, deoarece nu există terminații nervoase acolo, cu excepția durei mater.

În cele din urmă, profesorul Searle susține că atunci când un filosof spune că două persoane simt aceeași durere, ceea ce vrea să spună este că au același tip de durere, dar sunt două exemple distincte de durere (p. 116). „[C]a exemplu, durerea

pe care o experimentez ele sau eu există *doar în măsura în care este percepută de* un anumit subiect conștient" (p. 116). Fals. În primul rând, durerile nu sunt *percepute de* cei care le suferă. *A avea* o durere nu înseamnă *a percepe* o durere. „Am o durere în picior" nu descrie mai mult o relație între mine și un obiect numit „o durere" decât sintagma echivalentă „Mă doare piciorul". În al doilea rând, distincția lui Peirce între tip și exemplu se aplică inscripțiilor și depinde de convențiile ortografice. Se aplică durerilor nu mai mult decât s-ar putea aplica culorilor. Dacă două fotolii sunt maro, atunci există două fotolii de exact aceeași culoare, nu două exemple de același tip. Căci cum pot fi individualizate diferitele exemple? Evident, nu prin locație - deoarece locația distinge doar cele două fotolii, nu culoarea lor. Tot ce se poate spune este că primul presupus exemplu aparține primului fotoliu, iar al doilea celui de-al doilea fotoliu. Dar aceasta înseamnă individualizarea unei proprietăți prin referire la pseudo-proprietatea de a aparține substanței care o posedă - ca și cum proprietățile ar aparține substanțelor distinse prin legea lui Leibniz și ca și cum a fi proprietatea unei anumite substanțe ar fi o proprietate care distinge, de exemplu, culoarea *acestui* scaun de culoarea *celuilalt* . Și acest lucru este absurd. Cele două scaune au exact aceeași culoare. În mod similar, dacă două persoane au o durere de cap teribilă la tâmpla stângă, atunci ambele au exact aceeași durere de cap. Durerea de cap a lui A nu se distinge de durerea de cap a lui B prin faptul că aparține lui A, la fel cum culoarea maro a primului fotoliu nu se distinge de culoarea maro a celui de-al doilea fotoliu prin faptul că aparține primului. Distincția dintre identitatea calitativă și cea numerică nu se aplică culorilor sau durerilor , nici distincția lui Peirce între tipuri și exemplare .

ANTROPOLOGIE LINGVISTICĂ, AUTOANTROPOLOGIE, METAFORĂ ȘI EXTENSA DE UTILIZARE

Profesorul Dennett subliniază că examinarea utilizării cuvintelor necesită fie o

formă de antropologie, fie o „auto-antropologie”. În primul caz, trebuie descoperite uzul cuvintelor prin intermediul unor sondaje sociologice adecvate, cerându-le oamenilor să își ia în considerare intuițiile despre utilizarea corectă a unui cuvânt . O altă posibilitate este consultarea propriilor intuiții; dar s -ar putea dovedi că intuițiile cuiva diferă de cele ale altora. Profesorul Dennett afirmă că nu am consultat comunitatea neuroștiințifică pentru a le stabili intuițiile despre jargonul neuroștiințific al predicatelor psihologice (p. 86 și nota de subsol 15, p. 204), ci doar propriile noastre intuiții .

Aceasta este o greșeală. Un vorbitor competent al unei limbi nu trebuie să-și consulte intuițiile (bănuielile, conjecturile), la fel cum un matematician competent nu trebuie să-și consulte intuițiile despre tabla înmulțirii, nici un jucător de șah competent nu trebuie să-și consulte intuițiile despre mișcarea pieselor. Este un fapt empiric, pe care antropologii, istoricii limbajului etc. trebuie să-l stabilească, că un anumit cuvânt rostit sau scris este sau a fost folosit într-un anumit mod într-o anumită comunitate lingvistică. Nu este un fapt empiric că un cuvânt, care *are sensul pe care îl are,* are conexiunile conceptuale , compatibilitățile și incompatibilitățile pe care le are. Este un fapt empiric că vorbitorii de spaniolă folosesc cuvântul „negru” pentru a însemna ceea ce înseamnă, dar în măsura în care înseamnă ceea ce înseamnă, și anume, *această* culoare, nu este un fapt empiric că propozițiile „Negrul este mai închis decât albul”, „Negrul este mai asemănător cu griul decât albul”, „Nimic nu poate fi complet negru și *complet* alb în același timp” sunt adevărate. Sunt adevăruri conceptuale, care specifică o parte a rețelei conceptuale din care *„ negru”* este un nod. Incapacitatea de a recunoaște aceste adevăruri denotă o lipsă completă de înțelegere a sensului cuvântului. Un vorbitor competent este cel care stăpânește utilizarea expresiilor comune din limbă. Nu este intuiția lor că „negro” este *acea* culoare, că „yegua” este femela calului sau că „deambular ” înseamnă a rătăci. Nu este presimțirea lor că un bărbat este o ființă umană adultă de sex masculin. Și nu este presupunerea lor că dacă este ora zece este mai târziu de ora nouă sau că dacă ceva este complet negru nu este complet alb.

Deși vorbitorii competenți ai unei limbi sunt de acord asupra limbii pe care o vorbesc, abaterile de la uzul comun nu sunt, *în sine,* dăunătoare din punct de vedere filosofic. Astfel de abateri pot fi nimic mai mult decât un fragment dintr-un idiolect personal sau un sociolect special, o nouă extensie a unui termen sau însușirea unui termen inexistent pentru o nouă utilizare tehnică. Aceasta este ceea ce ne-a determinat să spunem că, dacă un vorbitor competent folosește expresii contrare uzului comun, atunci este foarte posibil ca...

> Cuvintele lor nu trebuie înțelese în sensul lor obișnuit. Poate că expresiile problematice au fost folosite într-un mod special și sunt, de fapt, doar omonime; sau au fost extensii *analogice* ale uzului comun, atât de răspândit în știință; sau au fost folosite metaforic sau figurativ. Dacă aceste căi de scăpare pot fi explorate, acuzația că neurologii au căzut pradă erorii mereologice este nejustificată.
>
> (PFN, p. 74)

Însă disponibilitatea acestor rute de evacuare nu este ceva ce poate fi considerat de la sine înțeles. Nici caracterul adecvat al acestei aplicări a termenului nu este o chestiune în care vorbitorul în cauză este autoritatea supremă. Deoarece, chiar dacă introduc o nouă utilizare sau își folosesc cuvintele figurativ , este necesar să vedem dacă o fac în mod coerent. Și este necesar să investigăm dacă o fac în mod consecvent sau dacă, dimpotrivă, trec în mod accidental între o utilizare nouă și una veche, făcând inferențe din prima care ar fi justificate doar de cea de-a doua. Din acest motiv am spus că

> Autoritatea supremă în această chestiune este *raționamentul propriu al vorbitorului.* Trebuie să analizăm concluziile pe care le trag din propriile - cuvinte, iar inferențele pe care le fac vor fi cele care vor demonstra dacă au folosit predicatul într-un sens nou sau l-au folosit greșit. Dacă trebuie să fie condamnați, trebuie să fie prin propriile cuvinte.

(PFN, p. 74)

Și apoi am continuat să arătăm că mulți neurocercetători importanți ar putea fi cu siguranță condamnați prin propriile lor cuvinte, tocmai pentru că trag inferențe din aplicarea lor particulară a vocabularului psihologic la creier, inferențe care pot fi făcute în mod inteligibil doar din aplicarea obișnuită a acelui vocabular la animal în întregime (PFN, cap. 3-8).

Dacă un neurocercetător aplică expresii psihologice sau semantice, cum ar fi „reprezentări" sau „hartă", creierului, fie folosește aceste expresii în sensul lor obișnuit, fie nu le folosește. Dacă nu sunt, atunci această utilizare neobișnuită poate implica 1) un sens derivativ, 2) o extensie analogică sau de altă natură a termenului mai vechi, 3) un simplu omonim sau *4)* un sens metaforic sau figurativ. Dacă termenii psihologici sunt aplicați creierului în sensul lor obișnuit, ceea ce se spune este ininteligibil. Nu știm ce înseamnă să spui că creierul gândește, se teme sau este rușinat. Criteriile constitutive pentru aplicarea, de către vorbitorii competenți ai limbii noastre, a unor astfel de expresii la animale *și* ființe umane - și anume, ceea ce spun *și* fac - nu pot fi îndeplinite de creier sau de părțile sale - nu există nimic asemănător cu un creier sau o parte a acestuia care face observații prudente, fuge de frică sau roșește de rușine. Nu înțelegem ce ar însemna pentru creier sau părți ale acestuia să gândească, să se teamă sau să decidă ceva mai bine decât înțelegem ce ar însemna pentru copaci să facă acest lucru. Dacă astfel de termeni sunt aplicați într-un sens nou, atunci utilizatorul ne datorează o explicație a acestui sens. Poate fi un sens derivativ, ca atunci când aplicăm termenul „sănătos" hranei sau exercițiilor fizice - o utilizare care necesită o explicație diferită de explicația adecvată a utilizării primare a termenului „sănătate" aplicat unei ființe vii. Poate fi o utilizare analogică, ca atunci când vorbim despre poalele muntelui sau ale paginii - astfel de analogii sunt de obicei evidente , dar în mod clar necesită o explicație parafratică foarte diferită de prototipul lor. Sau poate fi un omonim, cum ar fi „masă" în mecanica newtoniană, care necesită o explicație foarte diferită de cea a termenului „masă" în „o întreagă masă de oameni" sau „în flagrant delict".

Când neurocercetătorii vorbesc despre „reprezentare”, ei se referă , în cea mai mare parte, la un simplu omonim al termenului „reprezentare ” în sensul său simbolic și semantic. Această utilizare s-a dovedit nefericită , deoarece i-a determinat pe oamenii de știință și psihologi eminenți să cedeze confuziei de a folosi cuvântul atât pentru a se referi la o corelație cauzală sau concomitentă, cât și la o reprezentare simbolică. Dar numai în primul sens se poate vorbi despre reprezentări în creier. De aici și critica noastră la adresa lui David Marr (PFN, pp. 70, 76, 143–147). Se pare că utilizarea termenului „hartă” de către neurocercetători *a început* ca o extensie a ideii de „cartografiere ” , dar a fost curând confundată cu ideea de hartă. Nu este nimic greșit în a spune că un set de entități pe care pot fi cartografiați membrii unui alt set este o „hartă” a acestuia din urmă - deși această utilizare nu este nici necesară , nici complet clară. Dar inconsecvența apare atunci când se presupune că creierul ar putea folosi această „hartă” în modul în care cititorii unui atlas folosesc hărți. Nu este deloc clar ce se înțelege prin afirmația că creierul sau părțile sale știu, cred, gândesc, deduc și percep lucruri. Singura idee coerentă care ar putea fi sugerată prin aceasta este că acești termeni sunt aplicați creierului pentru a semnifica activitatea neuronală care se presupune că - corespunde cu cunoașterea, credința, gândirea, deducerea și percepția animalului . Dar din aceasta , afirmația (făcută de Crick, Sperry și Gazzaniga) că partea creierului care gândește comunică *ceea ce gândește* unei alte părți a creierului nu rezultă în niciun mod inteligibil . Deoarece, dacă gândirea umană are conținut (care oferă răspunsul la întrebarea „La ce te gândești ?”), nu se poate spune că activitatea neuronală are vreun fel de conținut.

S-ar putea sugera că, atunci când neurocercetătorii vorbesc despre *hărți* sau *descrieri simbolice din creier* și despre creier care cunoaște, gândește , decide, interpretează și așa mai departe, o fac metaforic.28 S -ar putea argumenta că acești termeni sunt, de fapt, metafore ascuțite, a căror potrivire a fost de mult stabilită în raport cu *computerele electronice,* care sunt descrise cu exactitate ca „respectând reguli”. Computerele *„au fost construite în mod deliberat* pentru a efectua «manipularea simbolurilor complexe conform regulilor». Într-adevăr , s-ar putea

crede că acest mod de a vorbi «nu mai este nici măcar metaforic, având în vedere cadrul teoretic și tehnologic extrem de dezvoltat în care are loc» [29] De asemenea, în utilizarea vocabularului psihologic comun, neurocercetătorii cognitivi „ bâjbâie de fapt în întuneric; metaforele sunt regula mai degrabă decât excepția". Dar aceasta face parte din progresul științific normal , iar în avansul său, neuroștiința a depășit metafora în unele cazuri, de exemplu, atribuind creierului „reprezentări asemănătoare propozițiilor" și „reprezentări asemănătoare hărților".

Toate acestea pot fi puse la îndoială. Este incorect să spunem că calculatoarele respectă reguli, la fel cum este incorect să spunem că planetele respectă legi. Legile lui Kepler descriu mișcarea orbitelor planetare, dar planetele în sine nu *respectă* aceste legi. Calculatoarele nu au fost construite „pentru a efectua manipularea simbolurilor conform regulilor", ci au fost construite pentru a produce rezultate care *corespund* manipulării *corecte* a simbolurilor conform regulilor . Calculatoarele nu pot *urma* o regulă, la fel cum nu poate face un calculator mecanic. O mașină poate efectua operațiuni care sunt conforme cu o regulă, cu condiția ca toate conexiunile cauzale încorporate în ea să funcționeze conform proiectării și presupunând că proiectarea asigură generarea unei regularități compatibile cu regula sau regulile alese. Dar pentru ca ceva să constituie un caz de respectare a unei reguli, simpla producere a unei regularități compatibile cu o regulă nu este suficientă. Se poate spune că o ființă urmează o regulă doar în contextul unei practici complexe care implică activități reale și potențiale de justificare a erorilor, recunoaștere și corectare a acestora prin referire la regulă, criticare a abaterilor de la regulă și, dacă este necesar, explicare a faptului că o acțiune este conformă cu regula și învățare a altora ce înseamnă să urmezi o regulă. Determinarea faptului că un act este *corect,* că este în conformitate cu regula, nu este o determinare cauzală, ci una logică. Altfel , ar trebui să acceptăm orice rezultate produc computerele. [30]

Fără îndoială, inginerii informaticieni folosesc acest tip de limbaj – inițial inofensiv, până când încep să-l aplice literal și presupun că, de fapt, computerele gândesc, că o fac mai bine și mai repede decât noi, că își amintesc cu adevărat și,

spre deosebire de noi, nu uită niciodată și că interpretează ceea ce tastăm – și uneori îl interpretează greșit, atribuind unui sens diferit de cel pe care intenționam să-l exprimăm. *Apoi,* modul de vorbire altfel inofensiv folosit de informaticieni încetează să mai fie o modalitate distractivă de a prescurta și devine o confuzie conceptuală potențial dăunătoare.

Este *posibil* ca cineva să spună că computerele sau creierul gândesc, calculează, raționează, deduc și formulează ipoteze doar metaforic. Metaforele nu explică - ele ilustrează un lucru prin intermediul altuia. O metaforă încetează să mai fie metaforă atunci când devine un clișeu, cum ar fi „o inimă frântă" sau „dintr-o singură lovitură", dar nu poate înceta să mai fie metaforă devenind literală. Ce ar însemna dacă ar fi *literal* adevărat că planetele respectă legile naturii? Se poate spune că o riglă de calcul, un calculator mecanic sau un computer calculează - figurativ vorbind. Dar ce ar însemna dacă ar calcula literal? Dacă „computerul își amintește (calculează, deduce etc.)" are sens perfect (non-figurativ) pentru inginerii de computere, este tocmai pentru că folosesc aceste expresii ca clișee. „Computerul calculează" nu înseamnă nimic mai mult decât „Computerul trece prin procesul electromecanic necesar pentru a produce rezultatele unui calcul fără niciun calcul", la fel cum „Te iubesc din toată inima" nu înseamnă nimic mai mult decât „Te iubesc cu adevărat".

Este clar că atunci când neurologii, în cazurile pe care le-am criticat, vorbesc despre hărți în creier și despre utilizarea de către creier a acestor hărți ca atare, nu o fac metaforic . Este evident că observația lui Blakemore conform căreia „neuroanatomiștii spun acum că creierul are *hărți,* despre care se crede că joacă un rol fundamental în reprezentarea și interpretarea lumii de către creier, *la fel cum o fac hărțile dintr-un atlas pentru cititorii săi" (sublinierea noastră)* nu este metaforică, deoarece nu este nimic metaforic în „Hărțile dintr-un atlas joacă un rol în reprezentarea lumii de către cititorii săi ". Mai mult, termenul „reprezentare" este folosit aici în mod manifest în sens simbolic, nu în sensul unei corelații cauzale. Nici afirmația lui JZ Young conform căreia creierul își folosește hărțile atunci când formulează ipoteze despre ceea ce este vizibil nu este metaforică.32 Deoarece

utilizarea unei *hărți în formularea unei ipoteze înseamnă* luarea unei caracteristici indicate de hartă ca *motiv* pentru ipoteză. Profesorul Dennett afirmă că creierul „ le folosește *ca* hărți" (p. 205, n. 20), iar în dezbaterea APA a afirmat că este o întrebare empirică, nu una filozofică, dacă creierul folosește „hărți retinotopice" *ca hărți,* dacă „oricare dintre operațiunile de recuperare a informațiilor definite în discuția despre ele exploatează caracteristicile hărților pe care le exploatăm atunci când facem acest lucru cu hărți ale lumii reale". Dar ar putea fi o întrebare empirică doar dacă ar avea sens ca creierul să folosească o hartă ca hartă. Cu toate acestea, pentru a *utiliza* o hartă ca hartă, trebuie să *existe* o hartă - *și* nu există una în creier - ; trebuie să știi cum să citești harta - dar creierul nu are ochi și nu poate citi -; trebuie să fii familiarizat cu convențiile de proiecție a hărților (de exemplu, proiecții cilindrice, conice, azimutale) - dar nu există convenții *de proiecție* pentru maparea caracteristicilor câmpului vizual pe activările neuronale din cortexul striat „vizual"; și trebuie să folosești harta pentru a-ți ghida comportamentul - excursiile sau navigațiile - o activitate pe care creierul nu o efectuează. O hartă nu trebuie confundată cu posibilitatea de mapare. Faptul că activarea celulelor retiniene poate fi mapată pe activarea celulelor din cortexul striat vizual nu demonstrează că există o hartă a ceea ce este vizibil în câmpul vizual al cortexului striat.

În final, am dori să corectăm o neînțelegere. Unii dintre criticii noștri presupun că încercăm să impunem o lege care interzice noi extensii ale expresiilor lingvistice. Profesorul Dennett a susținut în dezbaterea APA că am scoate în afara legii utilizarea „ codului genetic", deoarece insistăm că limbajul nu poate fi *codificat* în creier. Profesorul Churchland presupune că am exclude din principiu inovațiile conceptuale, cum ar fi afirmația lui Newton conform căreia Luna *cade* constant spre Pământ pe măsură ce se mișcă de-a lungul traiectoriei sale inerțiale. Toate acestea sunt o neînțelegere.

Nu interzicem nimic - doar semnalăm inconsecvențele conceptuale care apar în scrierile neuroștiințifice. Nu încercăm să împiedicăm pe nimeni să extindă utilizările limbajului într-un mod util din punct de vedere științific - doar căutăm să ne asigurăm că astfel de presupuse extinderi nu depășesc limitele sensului din

cauza incapacității de a specifica noua utilizare sau prin amestecarea necorespunzătoare a acesteia cu cea veche. Nu este nimic în neregulă cu a vorbi despre *poalele* muntelui, atâta timp cât nu ne întrebăm dacă muntele poartă încălțăminte. Nu este nimic în neregulă cu a vorbi despre *trecerea* timpului, atâta timp cât nu ne confundăm (așa cum se știe că a fost Sfântul Augustin) cu privire la modul de măsurare a acestuia. Nu era nimic în neregulă cu faptul că Newton spunea că Luna „cade", dar ar fi existat o problemă dacă ar fi întrebat ce o face să alunece. Nu era nimic în neregulă cu faptul că vorbea despre forțe care acționează asupra unui corp în spațiu, dar ar fi existat o problemă dacă ar fi speculat dacă forțele erau cele ale infanteriei sau cavaleriei. Nu este nimic în neregulă cu faptul că geneticienii vorbesc despre codul genetic. Dar ar exista o problemă dacă ar trage din existența sa concluzii care pot fi făcute doar din existența unor coduri literale. Pentru că, în realitate, codul genetic nu este un cod în sensul în care un cod este folosit pentru a codifica sau transmite o propoziție într-o limbă. Nu este nici măcar un cod într-un „sens discret ", cum ar fi o frază asupra căreia părinții au convenit, astfel încât copiii lor să nu știe despre ce vorbesc.

neurocercetătorii cognitivi utilizează vocabularul psihologic comun (și alți termeni, cum ar fi „reprezentare" și „hartă") atunci când specifică fenomenele pe care teoriile lor își propun să le explice și descriu termenii pe care îi folosesc pentru a face acest lucru. Deoarece, așa cum am clarificat , neurocercetătorii încearcă de obicei să explice percepția, cunoașterea, credința, memoria și luarea deciziilor umane referindu-se la părți ale creierului care percep, știu, cred, își amintesc și decid. Din acest motiv, am citat observații ale unor neurocercetători, psihologi și specialiști în științe cognitive de renume , cum ar fi următoarele:

> JZ Young: „Putem considera întreaga vedere ca o căutare continuă de răspunsuri la întrebările formulate de creier. Semnalele de pe retină constituie «mesaje» care conțin acele răspunsuri . Apoi, creierul folosește aceste informații pentru a construi o ipoteză adecvată despre ceea ce există acolo."
>
> *Programele creierului,* p. 119

C. Blakemore: „Creierul [are] hărți, despre care se crede că joacă un rol esențial în reprezentarea și interpretarea lumii de către creier , la fel cum o fac hărțile dintr-un atlas pentru cititorii săi."

„Înțelegerea imaginilor din creier", p. 265

G. Edelman: „Creierul «relaționează recursiv secvențele semantice cu cele fonologice și apoi generează corespondențe sintactice [...] tratând regulile care se dezvoltă în memorie ca obiecte ale manipulării conceptuale»."

Aer strălucitor, foc strălucitor (Harmondsworth, Penguin, 1994), p. 130

J. Frisby: „În creier trebuie să existe o descriere simbolică a lumii exterioare, o descriere formulată în simboluri care se referă la diversele aspecte ale lumii de care văzul ne face conștienți."

Văzând: Iluzie, Creier și Minte (Oxford, Oxford University Press, 1980), p. 8

E. Crick: „Când corpul calos este secționat, emisfera stângă vede doar jumătatea dreaptă a câmpului vizual [...] ambele emisfere pot auzi ceea ce se spune [...] se pare că o jumătate a creierului este complet inconștientă de ceea ce a văzut cealaltă jumătate."

Ipoteza uimitoare (Londra, Touchstone, 1995), p. 170

S. Zeki: „Capacitatea creierului de a dobândi cunoștințe, de a abstractiza și de a construi idealuri."

„Splendori și mizerii ale creierului", în *Tranzacții filosofice ale Societății Regale B* 354 (1999), p. 2.054

D. Marr: „Creierul nostru trebuie să fie cumva capabil să reprezinte [...] informații [...] Prin urmare, studiul vederii trebuie să includă o investigație a

naturii reprezentărilor interne prin care captăm aceste informații și le facem accesibile ca bază pentru decizii [...]"

O reprezentare este o schemă formală pentru descrierea [...] împreună cu regulile care specifică modul în care schema trebuie aplicată [...] [O schemă formală este] un set de simboluri cu reguli pentru combinarea lor. Prin urmare, o reprezentare nu este deloc o idee ciudată: cu toții folosim reprezentări în mod continuu .

Vision (San Francisco, Freeman, 1980), pp. 20 și următoarele.

Acestea nu sunt utilizări metaforice. Nu sunt extensii îndrăznețe ale unor termeni care introduc noi înțelesuri în scopuri teoretice. Sunt pur și simplu utilizări incorecte ale vocabularului psihologic (și semantic) comun - utilizări incorecte care duc la incoerență și diverse forme de contradicție - pe care le-am semnalat de la caz la caz. Nu este nimic ciudat în acest sens. Nu diferă, în principiu, de aplicările la fel de eronate ale aceluiași vocabular *la minte,* ca și cum ar fi mintea mea cea care știe, crede, gândește, percepe, simte durere, vrea și decide. Dar nu este așa; eu, ființa umană vie, sunt cea care face toate acestea. Prima eroare nu este mai puțin monumentală decât ultima (și venerabila) eroare și este răspândită printre neurocercetătorii cognitivi, uneori în detrimentul experimentelor pe care le concep, de obicei în detrimentul teoretizării rezultatelor acelor experimente și foarte des în detrimentul explicării funcțiilor cognitive animale și umane prin referire la structurile și operațiile neuronale care le fac posibile.

CALIA

În expunerea noastră despre conștiință (PFN, cap. 9-12), am susținut că caracterizarea domeniului mentalului prin referire la „sentimentul calitativ" al experienței este eronată (PFN, cap. 10). Dar, cu permisiunea profesorului Searle (p. 99 și următoarele), nu negăm existența *qualia* susținând că, dacă acestea ar

exista, ar exista în creier. *Si, per impossibile,* Dacă toate atributele psihologice ar fi caracterizate prin „sentimentul lor calitativ", ele ar fi tot atribute ale ființelor umane, nu ale creierului.

quale ar trebui să fie este „sentimentul calitativ al unei experiențe " (Chalmers), [33] sau este ceva de genul „roșeața roșului sau durerea durerii" (Crick). [34] Qualia sunt „calitățile senzoriale simple găsite în albastrul cerului sau în sunetul unei *note* " (Damasio) [35] sau „ceea ce simte cineva atunci când vede, aude sau miroase, ceea ce simte atunci când experimentează durere" (Block). [36] Potrivit profesorului Searle, stările conștiente sunt „calitative în sensul că, pentru fiecare stare conștientă [...] există ceva care este calitativ similar cu a fi în acea stare". [37] Potrivit lui Nagel, pentru fiecare experiență conștientă „există ceva care este similar cu a o avea pentru organism". [38] Aceste explicații diverse *nu sunt echivalente și nu este clar că* din ele rezultă o explicație coerentă .

Profesorul Searle observă că într-o durere, o mâncărime sau o gâdilare există o senzație calitativă. Suntem de acord, în sensul următor: senzațiile, am subliniat (PFN, p. 124), posedă calități fenomenale (de exemplu, cele de arsură, usturime, mâncărime, rontăit, înțepături); sunt asociate cu înclinația către anumite comportamente (suflat, scărpinat, frecat sau râs); și au grade de intensitate care pot crește sau scădea.

Cu toate acestea, am observat că, atunci când vine vorba de percepție, este dificil să caracterizăm ce se înțelege prin „natura calitativă a experienței". Specificarea a ceea ce vedem sau mirosim sau, în cazul halucinațiilor, a ceea ce credem că vedem sau mirosim, necesită specificarea unui obiect. Experiențele vizuale sau olfactive și halucinațiile corespunzătoare acestora sunt individualizate prin ceea ce reprezintă experiențe sau halucinații. A vedea un stâlp de iluminat este diferit de a vedea o cutie poștală, a mirosi un liliac este diferit de a mirosi un trandafir și la fel sunt și experiențele halucinatorii corespunzătoare, care sunt descrise în funcție de similaritatea pe care o prezintă subiectului cu adevăratul lor corespondent perceptiv.39

Trandafirii cu siguranță nu miros a liliac - felul în care mirosul trandafirilor

este diferit de felul în care mirosul liliacului. A mirosi trandafirii este foarte diferit de a mirosi liliacul. Dar aspectul calitativ al mirosirii trandafirilor nu miroase a trandafiri și nici aspectul calitativ al mirosirii liliacului nu miroase a liliac. Mirosirea oricăreia dintre acestea poate fi la fel de plăcută - caz în care aspectul calitativ al mirosului poate fi exact același, chiar dacă ceea ce se miroase este diferit. Credem că profesorul Searle confundă modul în care sunt percepute mirosurile cu modul în care este perceput mirosul.

În mod normal, a vedea un stâlp de iluminat nu evocă sentimente. Dacă întrebi „Ce ai simțit când l-ai văzut?", singurul tip de răspuns pe care îl vei primi este ceva de genul „Nu am simțit nimic anume, nici plăcut, nici neplăcut, nici incitant, *nici* plictisitor". *Aceste* epitete - „plăcut", „neplăcut", „incitant ", „plictisitor" - *sunt înțelese* pe bună dreptate ca descrieri ale „caracterului calitativ al experienței". În acest sens, *multe* experiențe perceptive nu au niciun caracter calitativ. *Niciuna nu este* individualizată prin senzația sa calitativă - sunt individualizate prin obiectul său. Și dacă vorbim despre o halucinație, a spune că stâlpul de iluminat din halucinația noastră era negru este totuși o descriere a obiectului experienței - „obiectul său intențional" în jargonul lui Brentano (pe care îl folosește profesorul Searle). Pe de altă parte, calitatea experienței halucinatorii este probabil *copleșitoare.*

Contrar a ceea ce sugerează profesorul Searle, nu susținem că „dacă *qualia* nu sunt definite ca o chestiune de agreabilitate sau dezagreabilitate, atunci experiența trebuie individualizată prin obiectul său" (p. 115). Teza noastră a fost că individualizăm *experiențele* și halucinațiile prin obiectele lor - care sunt specificate de răspunsul la întrebarea „ Despre ce a fost experiența (sau halucinația) dumneavoastră?" - [40.] Desigur , obiectul nu trebuie să fie cauza, așa cum este evident în cazul halucinațiilor. Dar, am insistat, caracterul calitativ al experienței nu trebuie confundat cu calitățile obiectului experienței. Faptul că ceea ce vede cineva atunci când privește un măr roșu este roșu și rotund nu implică faptul că a avut o experiență vizuală roșie și rotundă. Faptul că ceea ce crede că vede cineva atunci când halucinează un măr roșu este roșu și rotund nu implică faptul că a avut

de fapt o halucinație vizuală roșie și rotundă. „Ce ați văzut (sau halucinat)?" este o întrebare, iar „Cum a fost să vezi ce ai văzut (sau să halucinezi ce ai halucinat)?" este o alta. Experiențele perceptive nu sunt individualizate prin natura lor calitativă. Acestea sunt adevăruri simple, dar par să fi fost trecute cu vederea.

CREIERE ÎN INTERIORUL UNUI CRANIU

Profesorul Searle spune că ființele umane sunt „creiere într-un corp" (pp. 120 și următoarele). Din punctul său de vedere, motivul pentru care putem spune atât „Cântăresc 75 de kilograme", cât și „Corpul meu cântărește 75 de kilograme" este că ceea ce mă face să cântăresc 75 de kilograme este faptul că asta cântărește corpul meu. Dar, aparent, strict vorbind, nu sunt nimic mai mult decât un creier într-un corp (un craniu). *Am* un corp și sunt *în* craniul *corpului* meu. Aceasta este o versiune materialistă a cartesianismului. Unul dintre principalele motive pentru care am scris cartea noastră provine din convingerea fermă că neurologii contemporani, și mulți filosofi, se află încă sub umbra lungă și întunecată a lui Descartes. Deși resping substanța imaterială a minții carteziene, ei transferă atributele minții carteziene creierului uman, lăsând intactă întreaga structură a concepției carteziene a relației minte-corp. Ceea ce susținem este ca neurocercetătorii, și chiar filosofii, să lase în urmă tărâmurile carteziene întunecate și să caute lumina soarelui aristotelică, unde apare o viziune mult mai clară.

Da, *oricât de imposibil,* Dacă aș fi un creier în interiorul unui corp, atunci aș avea un corp - așa cum mintea cartesiană are un corp. Dar nu aș *avea* un creier, deoarece creierele nu au creier. Și, în realitate, corpul meu nu ar cântări 75 de kilograme, ci 75 de kilograme minus cele 1,5 kilograme care, la propriu vorbind, este cât aș cântări *eu* . Și nu aș avea 1,75 metri înălțime, ci doar 0,18 metri. Fără îndoială, profesorul Searle mă va asigura că sunt creierul meu în interiorul corpului meu - creierul meu *împreună* cu corpul meu. Dar asta nu ne duce nicăieri. Căci creierul meu împreună cu corpul meu fără creier, luat într-un sens, este pur și simplu cadavrul meu ; luat în alt sens, este pur și simplu *corpul meu.* Dar eu nu sunt corpul meu, nu corpul pe care *îl am.* Firește, eu sunt *un* corp - ființa umană

vie din fața cititorului, un anumit tip de ființă spațio-temporală persistentă, simțitoare, care posedă inteligență și voință și, prin urmare, o persoană. Dar eu nu sunt mai mult *corpul meu* decât *mintea mea.* Nici nu sunt un creier în interiorul unui corp. Este o greșeală să presupunem că ființele umane sunt „în interiorul corpurilor" – această idee aparține tradiției platonice, augustiniene și carteziene, care ar trebui respinsă. Ar fi mult mai bine să spunem, împreună cu Aristotel, că ființele umane sunt *creaturi animate (empsuchi),* animale înzestrate cu capacități care le conferă, în forma lor naturală de viață, statutul de persoane .

CERCETAREA ÎN NEUROȘTIINȚE

Criticii noștri susțin că cercetările noastre sunt irelevante pentru neuroștiințe sau, mai rău, că sfaturile noastre, dacă ar fi urmate, ar fi pur și simplu dăunătoare. Profesorul Dennett susține că refuzul nostru de a atribui atribute psihologice (chiar și într-un sens limitat) oricărui lucru decât unui animal întreg este înapoiat și neștiințific, în contrast cu beneficiile științifice ale „atitudinii intenționale" pe care o susține. În opinia sa, „licența poetică oferită de atitudinea intențională facilitează" explicarea modului în care funcționarea părților contribuie la -comportamentul animalului (p. 89).

În primul rând, observăm că *libertatea poetică* este ceva permis poeților în scopul poeziei, nu în scopul acurateței empirice și al puterii explicative. În al doilea rând, atribuirea abilităților cognitive unor părți ale creierului oferă doar aparența unei explicații atunci când este necesară o explicație adevărată și, prin urmare, împiedică procesul științific. Sperry și Gazzaniga susțin că, în cazurile de comisurotomie, comportamentul ciudat al subiecților experimentali expuși la reprezentări ale obiectelor este *explicat* prin faptul că o emisferă a creierului nu este conștientă de ceea ce cealaltă jumătate poate vedea. §§Se presupune că

§§Într-o serie celebră de experimente, Gazzaniga, Bogen și Sperry (1962) au prezentat pacienților cu comisurotomie stimuli vizuali și tactili direcționați alternativ

emisferele creierului știu lucrurile și sunt capabile să le explice și că, din cauza secționării corpului calos, emisfera dreaptă nu poate comunica celei stângi ceea ce vede, așa că aceasta din urmă trebuie să genereze propria interpretare a motivului pentru care mâna stângă face ceea ce face. [Departe] de a explica fenomenele, aceasta este doar o redescriere a acestora în termeni confuzi, ascunzând astfel absența unei explicații substanțiale. Disocierea funcțiilor asociate în mod normal este explicată parțial prin secționarea corpului calos și prin localizarea funcțiilor în ambele emisfere. *Acest lucru* este acum perfect cunoscut, dar explicația disponibilă în prezent nu merge mai departe. Este o iluzie să presupunem că nu se adaugă absolut nimic acestei explicații prin atribuirea a ceva asemănător („ceva asemănător" sau altceva).
cunoaștere, percepție și înțelegere lingvistică către emisferele creierului.

Profesorul Searle susține că, dacă explicația noastră privind structurile conceptuale utilizate ar fi corectă, întrebările esențiale ale cercetării neurobiologice ar fi respinse ca fiind lipsite de sens. Din acest motiv, el subliniază că „întrebarea centrală a vederii, modul în care procesele neurobiologice [...] provoacă experiențe vizuale conștiente, nu ar putea fi investigată de nimeni care ar accepta concepția [noastră]". Concepția noastră, afirmă el, „ar putea avea - consecințe științifice potențial dezastruoase" (p. 124).

Cercetarea în neurobiologia vederii este o cercetare a structurilor neuronale necesare cauzal pentru ca un animal să vadă și a proceselor specifice implicate în

către una dintre cele două emisfere, care erau „neconectate" din cauza absenței corpului calos. Stimulii vizuali prezentați câmpului vizual drept (conectat la emisfera stângă) au fost raportați corect de către subiecți, dar când aceiași stimuli au fost prezentați câmpului vizual stâng (conectat la emisfera dreaptă), subiecții au susținut că nu au văzut nimic. Dacă li s-a cerut apoi să ridice un obiect folosind mâna stângă (controlată de aceeași emisferă dreaptă), au făcut-o fără dificultate, dar dacă li s-a cerut apoi ce țin în mână, subiecții nu au putut răspunde . Vezi Gazzaniga, MS, Bogen, JE și Sperry, RW (1962) „Câteva efecte funcționale ale secționării comisurilor cerebrale la om", în *Proceedings of the National Academy of Sciences,* 48, partea a 2-a, p. 1.765. *(Nr. titlului)*

vedere. Respingerea ideii că experiențele vizuale au loc în creier sau că acestea sunt caracterizate de *qualia* afectează acest program de cercetare în neuroștiințe doar în măsura în care evită întrebări inutile care ar putea rămâne fără răspuns. Dăm numeroase exemple: problema integrării (Crick, Kandel și Wurtz), explicația recunoașterii prin referire la compararea șabloanelor și imaginilor (Marr) sau sugestia că percepțiile sunt ipoteze ale creierului, concluzii trase din inferențele inconștiente pe care le face (Helmholtz, Gregory și Blakemore). Concepția noastră că animalul, nu creierul, vede sau are experiențe vizuale și concepția profesorului Searle că creierul , nu animalul, sunt teze conceptuale, nu empirice. Subiectul este, fără îndoială, important din toate motivele menționate, dar ar trebui să fie clar că ceea ce am spus nu *împiedică* cercetarea empirică asupra proceselor neuronale care stau la baza vederii. Dimpotrivă, ea orientează descrierea rezultatelor unor astfel de cercetări spre calea sensului.

În general, criticile conceptuale din cartea noastră nu fac altceva decât să îndepărteze straturile de confuzie conceptuală din studiile neuroștiințifice și să clarifice cadrele conceptuale pe care le presupun. Acest lucru nu poate împiedica progresul neuroștiinței. De fapt, ar trebui să îl faciliteze prin excluderea întrebărilor fără sens, interzicerea experimentelor greșit direcționate și eliminarea rezultatelor experimentale interpretate greșit. [42]

EPILOG

MAXWELL BENNETT

L-am întâlnit pe Sir John Eccles în 1962, când îmi terminam studiile universitare de inginerie electrică. În anul următor, a câștigat Premiul Nobel pentru lucrările sale asupra transmiterii chimice la sinapsele din măduva spinării și creier. M-a întrebat ce fac și i-am spus: „Inginerie electrică", la care mi-a răspuns: „Excelent, trebuie să vii cu mine, pentru că fiecare laborator neurofiziologic de primă clasă are nevoie de un sudor bun". Cred că fiecare laborator de neuroștiințe cognitive de primă clasă are nevoie astăzi de un bun filosof critic și analitic. Dialogul despre obiectivele și realizările neuroștiinței cognitive reprodus în această carte , care a avut loc la întâlnirea APA din 2005 de la New York, confirmă această convingere.

Sarcina neuroștiinței este de a înțelege funcționarea sistemului nervos, contribuind astfel la conceperea strategiilor pentru a atenua povara teribilă a bolilor precum demența și schizofrenia asupra umanității. În îndeplinirea misiunii lor, neurocercetătorii clarifică, de asemenea, acele mecanisme cerebrale care trebuie să funcționeze normal pentru ca noi să ne putem exersa facultățile psihologice, cum ar fi percepția și memoria .

Această concepție despre neuroștiință se opune celei care susține că neuroștiința are un singur scop general: înțelegerea conștiinței.1 *Este* interesant să analizăm în detaliu această sugestie, ca exemplu al necesității unei analize filosofice critice a neuroștiințelor de tipul celei pe care am propus-o mai sus. Peter Hacker și cu mine dedicăm peste o sută de pagini în cartea noastră temei *conștiinței* . La începutul analizei noastre, afirmăm că „un prim pas către claritate este să distingem conștiința *tranzitivă* de conștiința intranzitivă . *Conștiința tranzitivă* este capacitatea de a fi conștient *de* un lucru sau altul sau de a fi conștient *că* un lucru sau altul este într-un fel sau altul. Conștiința intranzitivă, pe de altă parte, nu are niciun obiect. Este capacitatea de a fi conștient sau treaz, spre deosebire de a fi inconștient sau adormit" *{Fundamentele filozofice ale neuroștiinței }*. (denumită în continuare PFN, p. 244). Pierderea conștiinței intranzitive, de exemplu, atunci când cineva doarme, este inconștient sau sub anestezie, face obiectul unei bogate literaturi neuroștiințifice. Pe de altă parte, există multă confuzie în literatura neuroștiințifică atunci când vine vorba de studiul diferitelor forme de conștiință tranzitivă. Aceste forme includ conștiința perceptivă, somatică, kinestezică și afectivă, conștientizarea propriilor motivații, conștiința reflexivă, conștientizarea propriilor acțiuni și conștiința de sine (PFN, pp. 248-252). Unele dintre aceste forme de conștiință tranzitivă sunt atenționale. De exemplu, conștientizarea perceptivă a ceva implică faptul că lucrul de care este conștient atrage atenția într-un mod susținut. Este o greșeală să presupunem că percepția a ceva este, în măsura în care este percepție, o formă de conștiință tranzitivă sau chiar că implică a fi conștient de ceea ce percepe.

Cercetarea în neuroștiințe s-a dedicat în mare măsură unei singure forme a ceea ce este interpretat (sau interpretat greșit) drept conștiință perceptivă tranzitivă, în special aceleia care...

Acest lucru joacă un rol în percepția vizuală, în special în fenomenul rivalității binoculare. În timpul acestei rivalități, observatorul surprinde două imagini incongruente, fiecare afectând un ochi diferit, dar percepe o singură imagine la un moment dat. Imaginile dominante perceptiv și cele suprimate alternează la fiecare câteva secunde. Munca experimentală a lui Logothetis și colegilor săi (Leopold și Logothetis 1999; Blake și Logothetis 2002) ajută la înțelegerea acestui lucru. Folosind tehnici de condiționare operantă, au antrenat maimuțe să apese o pârghie pentru a indica care dintre cele două imagini monoculare concurente este dominantă la un moment dat (vezi Figura 8). Vârfurile activității potențialului de acțiune, înregistrate în celulele individuale ale cortexului vizual al maimuței, pot fi corelate cu răspunsul perceptiv al animalului la apăsarea pârghiei. Acest lucru permite identificarea regiunilor corticale în care activitatea neuronală corespunde experienței perceptive. Caseta din stânga Figurii 8 arată numărul de vârfuri de activitate electrică înregistrate într-un...

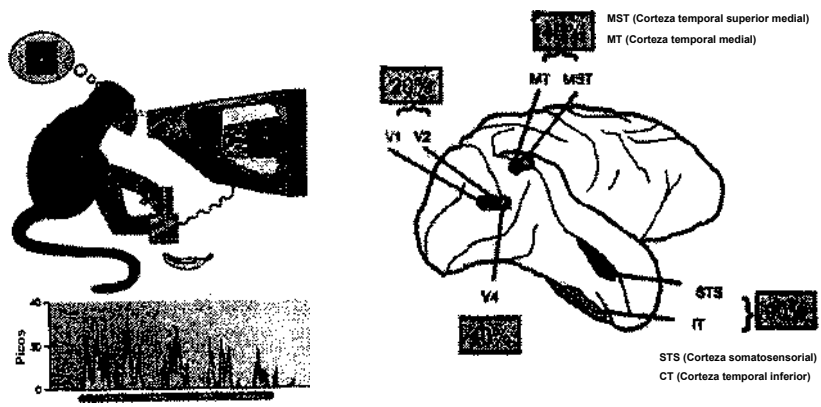


FIGURA 8: Tehnici de condiționare operantă pentru a stabili zonele cortexului cu neuroni care, în timpul prezentării imaginilor monoculare „rivale", descarcă vârfuri de potențiale de acțiune corelate cu „rapoartele perceptive" ale maimuțelor (adaptat după Blake și Logothetis, 2002; Leopold și Logothetis, 1999). Bara de sub axa x indică percepția alternativă a celor două imagini, în corelație clară cu perioadele de activitate maximă. Figura 8 (dreapta) prezintă regiunile creierului care conțin neuroni excitați a căror activitate se corelează cu percepția vizuală a maimuței. Procentul de neuroni legați de percepție crește în centrii vizuali „superiori", adică în acele zone ale cortexului mai îndepărtate de inputul din talamus. Doar o mică fracțiune dintre neuronii sensibili la stimuli

vizuali din zonele corticale V1 (cu conexiuni directe cu talamusul) și V2, primele zone care primesc input talamic, au răspuns în funcție de alternanțele rivalității binoculare. În schimb, procentul a fost mai mare în zonele mai îndepărtate de inputul talamic , în special V4, cortexul temporal medial (MT) și cortexul temporal medial superior (MST). Activitatea aproape tuturor neuronilor sensibili la stimuli vizuali din cortexul temporal inferior (IT) și somatosenzorial (STS) a coincis în mare măsură cu starea perceptivă a animalului. Această modulare a neuronilor corticali contrastează cu activitatea neuronilor non-corticali din nucleul geniculat lateral, care au conexiuni directe cu retina. Aceștia nu prezintă nicio modulare în timpul rivalității binoculare.

Astfel, există un set distribuit de neuroni în cortex care sunt activați în coordonare cu reacțiile perceptive ale animalului, deși mai puțini neuroni sunt activați în zonele vizuale primare (VI) decât în zonele vizuale superioare. Lumer și colegii săi, utilizând imagistica RMN funcțională la subiecți umani, au determinat distribuția spațială a neuronilor corticali care sunt activi în timpul acestor reacții la alternanța binoculară. Studiile lor arată că neuronii a căror activitate fluctuează în coordonare cu fluctuațiile raportate ale experienței perceptive sunt distribuiți pe scară largă în cortex, inclusiv cortexul prefrontal lateral și centrii vizuali superiori (Lumer, Friston și Rees, 1998). Edelman și colegii săi au observat același lucru , utilizând magnetoencefalografia (MEG) pentru a înregistra activitatea corticală a subiecților umani (Tononi și Edelman, 1998; Srinivasan și colab., 1999). Acest tip de activitate distribuită în timpul percepției pare să fie ceea ce l-a determinat pe profesorul Searle să sugereze că există un „câmp concomitent”.

• • 2
știință».

Spre deosebire de accentul pus pe distribuția activității corticale în timpul rivalității binoculare, alți autori subliniază că doar anumite clase de neuroni din centrele vizuale superioare pot fi considerate candidați pentru corelația neuronală a conștiinței (NCC) pentru acest fenomen. De exemplu, Crick și Koch (2003) subliniază că doar centrele vizuale superioare, cum ar fi STS și IT, posedă neuroni sensibili la stimuli vizuali care, de obicei (90%), reacționează în coordonare cu

alternanțele rivalității binoculare. Acest lucru îi determină pe Crick și Koch să examineze detaliile „arborizărilor dendritice ale diferitelor tipuri de neuroni din girusul temporal inferior (IT) al macacului care se proiectează spre cortexul prefrontal, în apropierea șanțului principal” (Fig. 9; mai sus, evidențiată cu gri). Ei observă că „doar un singur tip de celulă are dendrite apicale care ating stratul 1”. Apoi, ei întreabă: „Ce este special la activitatea care depășește pragul conștiinței ? Ar putea fi declanșarea unor tipuri speciale de neuroni, cum ar fi celulele piramidale care își proiectează axonii spre partea din față a creierului” (Fig. 9). Adică, acum avem că receptorii celulelor neuronale (NCR) ai percepției în timpul rivalității binoculare sunt identificați cu un anumit tip de neuron din cortexul temporal inferior, o idee similară cu cea a unei „celule pontife” sau „celule cardinale” a

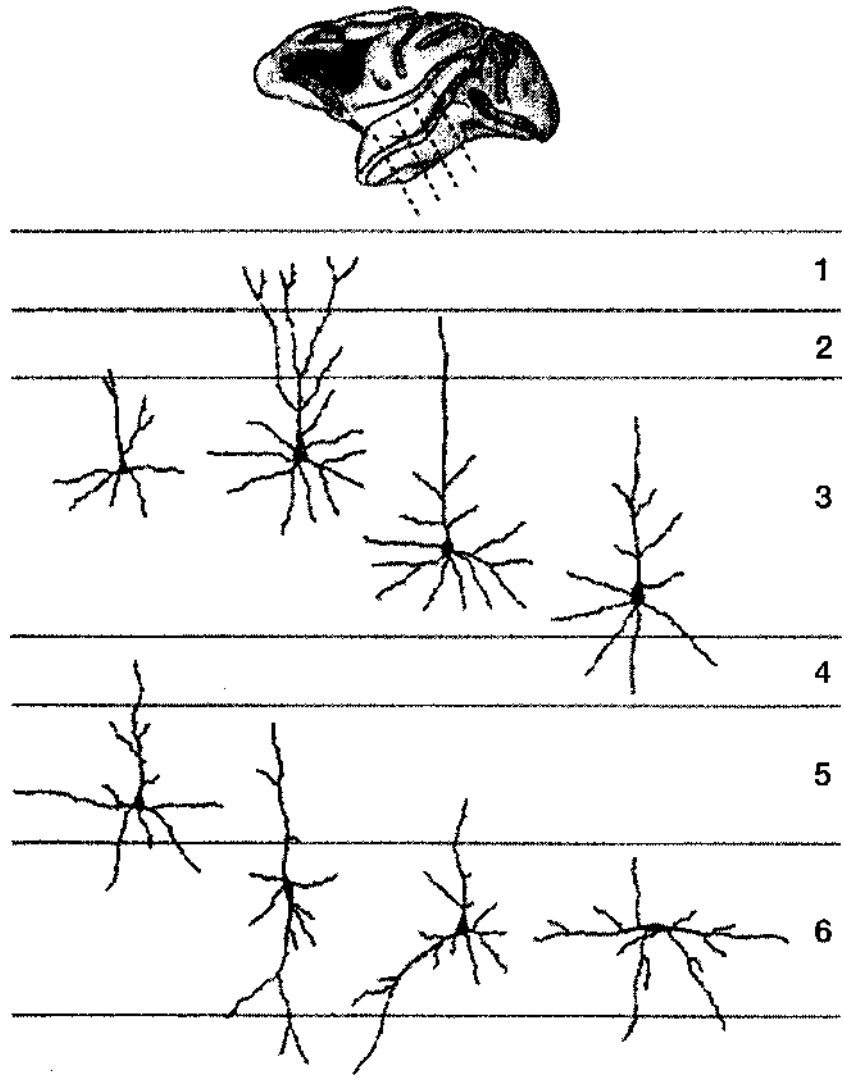


FIGURA 9: Colecție de neuroni din cortexul temporal inferior (IT în Figura 8) care se proiectează către o porțiune limitată a cortexului prefrontal (zona punctată a creierului reprezentată în partea de sus; preluat din Crick și Koch, 2003, din De Lima, Voight și Morrison, 1990).

Barlow (1997). ***Acest lucru îl conduce pe Koch (2004), în cartea sa **The Quest*

***Aceasta face aluzie la o metaforă a lui Sherrington (1941) care sugerează că sistemele perceptive corticale nu pot fi organizate în ierarhii statice cu un singur neuron „pontif" la vârf. Barlow a reexaminat critic această idee pentru a pleda pentru

for Consciousness,* foarte admirată de venerabilii filozofi, să descrie, folosind Figura 10, ceea ce se presupune că neuroștiința a dezvăluit despre CNC în timpul percepției vizuale . Aceasta este o concepție cripto-cartezianăa relației dintre atributele noastre psihologice *și* funcția creierului . Este o idee înrădăcinată în credința falsă, pe care o subliniem în cartea noastră (PFN, Capitolul 10), că esența conștiinței este asocierea sa cu *qualia și* că, prin urmare, aceste caracteristici calitative inefabile ale experienței pot fi descoperite ca fiind cauzate de celule cardinale sau pontificale.

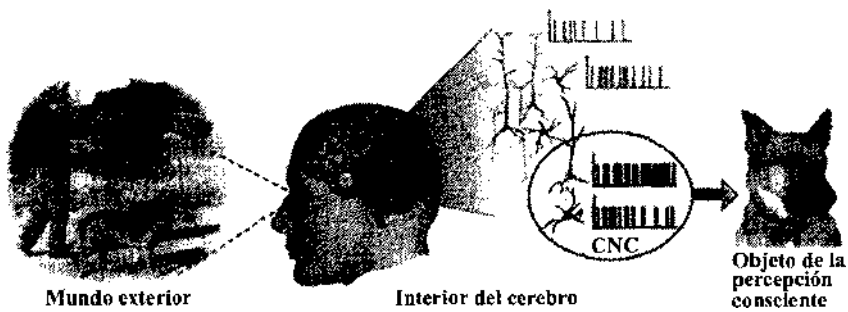


FIGURA 10: „CNC-urile reprezintă setul minim de evenimente neuronale — aici potențiale de acțiune sincronizate ale neuronilor piramidali neocorticali (ale zonei IT) — suficiente pentru o anumită percepție conștientă" (extras din Koch, 2004),

Cred că această scurtă expunere a cercetării în neuroștiințe și interpretarea percepțiilor vizuale în timpul rivalității binoculare relevă o nevoie stringentă de clarificare critică din partea filosofilor analitici. Peter Hacker și cu mine sugerăm că rezultatele acestei cercetări, fie că sunt interpretate în termeni de „câmp de conștiință" sau „ celule cardinale", nu contribuie la înțelegerea conștiinței perceptive tranzitive. În cel mai bun caz, ele contribuie la identificarea unora dintre corelațiile neuronale ale percepției vizuale în condiții de rivalitate binoculară. Însă conștiința perceptivă tranzitivă implică faptul că un obiect din câmpul vizual ne

„asamblaje cardinale", agregări dinamice de celule în vârful ierarhiilor la fel de statice, subliniind importanța activității neuronale specifice pentru percepție. Vezi Sherrington, CS, *Man on his nature ,* Cambridge, Cambridge University Press, 1941 (traducere în spaniolă: *El hombre en su naturaleza,* Madrid, Alhambra, 1947).

atrage atenția într-un mod susținut. Pentru a descoperi corelațiile neuronale ale conștiinței perceptive tranzitive, un studiu în neuroștiințe trebuie să se refere la corelațiile neuronale ale atenției captive față de ceea ce percepem, nu la percepția în sine. Căci a percepe un obiect nu este același lucru cu a fi conștient de obiectul perceput . Cineva poate percepe un obiect X fără a fi conștient de el, fie pentru că îl identifică în mod eronat ca fiind Y, fie pentru că atenția sa nu este atrasă în mod constant către el, ceea ce ar putea fi pentru că nici măcar nu îl percepe sau pentru că se *concentrează intenționat asupra lui.* Conștientizarea perceptivă tranzitivă - este o formă de *receptivitate* cognitivă (PFN, pp. 253-260). Conștientizarea afectivă, conștientizarea propriilor motivații, conștientizarea reflexivă a propriilor acțiuni și conștientizarea de sine necesită forme de analiză care diferă în aspecte importante. În *Fundamentele filosofice ale neuroștiinței* și în noua noastră carte *, Istoria neuroștiinței cognitive: o analiză conceptuală ,* În continuare, vor fi incluse multe alte exemple care demonstrează necesitatea clarificării și care cuprind - întreaga gamă de cercetări neuroștiințifice privind corelațiile neuronale ale atributelor noastre psihologice.

Dacă filosofii devin acoliți ai demersului neuroștiințific , nu vor face decât să amplifica divinizarea neuroștiinței pe care am menționat-o în introducere. Ceea ce are nevoie disciplina nu sunt recenzii laudative în *New York Review of Books.* În ceea ce privește lucrările neurocercetătorilor aparent inconștienți de dificultățile conceptuale inerente ideilor lor, ceea ce avem nevoie este o critică filosofică revelatoare, care să poată ghida cercetări neuroștiințifice fructuoase asupra capacităților noastre psihologice și a exercitării acestora. Aceasta, în opinia mea, este o sarcină importantă pentru noile generații de filosofi.

REFERINȚE BIBLIOGRAFICE

Barlow, HB și WR Levick, „Mecanismul unităților selective direcționale din retina iepurelui ”, în *Journal of Physiology,* nr. 178, 1965, pag. 477-504.

Bennett, MR, *Istoria sinapsei,* Amsterdam, Academia Harwood , 2001.

Bennett, MR, L. Farnell și WG Gibson, „Un model cantitativ al transmiterii joncționale purinergice a undelor de calciu în rețelele de astrocite", în *Biophysics Journal,* nr. 89, 2005, pag. 2.235-2.250.

Bennett, MR, WG Gibson și J. Robinson, „Dinamica rețelei de memorie autoasociativă a neuronilor piramidali CA3 în campusul hipopotamului ", în *Philosophical Transactions of the Royal Society of London B Biological Sciences,* nr. 343, 1994, pag. 167-187.

Bennett, MR și PMS Hacker, *Fundamentele filozofice ale neuroștiinței ,* Oxford, Blackwell, 2003.

Bennett, MR și AG Pettigrew, „Formarea sinapselor neuromusculare", în *Cold Spring Harbor Symposium in Quantitative Biology,* nr. 40, 1975, pag. 409-424.

Blakemore, C., *Mecanismele minții,* Cambridge, Cambridge University Press, 1997.

Bridley, GS, „Clasificarea sinapselor modificabile și utilizarea lor în modele de condiționare", în *Proceedings of the Royal Society London B,* nr. 168, 1967, pag. 361-376.

—, «Modele de rețele nervoase de dimensiuni plauzibile care îndeplinesc multe sarcini simple de învățare», în *Proceedings of the Royal Society London B,* nr. 174, 1969, pag. 173-191.

Brodman, K., *Localizare comparativă: Teoria cortexului cerebral în principiile sale prezentate pe baza structurii celulare,* Leipzig, Barth, 1909.

Cornell-Bell, AH, SM Finkbeiner, MS Cooper și S. J. Smith, „Glutamatul induce unde de calciu în astrocitele cultivate: semnalizare glială pe termen lung ", în *Science,* nr. 247, 1990, pag. 470-473.

Fried, SI, TA Munch și FS Werblin, „Mecanisme și circuite care stau la baza selectivității direcționale în retină", în *Nature,* nr. 420, 2002, pag. 411-414.

—, «Selectivitatea direcțională se formează la niveluri multiple prin inhibarea decalată lateral în retina iepurelui», în *Neuron,* nr. 46, 2005, pag. 117-127.

Gazzaniga, MS, RB Ivry și GR Mangun, *Neuroștiința cognitivă: Biologia minții, ediția* a 2 -a , Nueva York, Norton, 2002.

Giannikopoulos, DV și U. T. Eysel, «Dinamica și specificitatea reorganizării hărții corticale modifică leziunile retiniene», în *Proceedings of the National Academy of Sciences, SUA, nr.* 103, 2006, pag. 1.085-1.010.

Gibson, WG, J. Robinson și M. R. Bennett, „Secreția probabilistică de cuante în sistemul nervos central: Controlul sinaptic al celulelor granulare asupra separării tiparelor și reglării activității", în *Philosophical Transactions of the Royal Society of London B Biological Sciences,* nr. 332, 1991, pag. 199-220.

Gilbert, CD, «Dinamica corticală la adulți», în *Physiology Review,* nr. 78, 1998, pag. 467-485.

Gilbert, CD, JA Hirsch și TN Wiesel, «Interacțiuni laterale în cortexul vizual», în *Cold Spring Harbor Symposium in Quantitative Biology ,* nr. 55, 1990, pag. 663-677.

Gilbert, CD și T. N. Wiesel, «Dinamica câmpului receptiv în cortexul vizual primar la adult», în *Nature,* nr. 356, 1992, pag. 150-152.

Hubei, DH, TN Wiesel și S. LeVay, „Arhitectura funcțională a zonei 17 la maimuțele macaci normale și cu deficiențe monoculare", în *Cold Spring Harbor Symposium in Quantitative Biology,* nr. 40, 1976, pag. 581-589.

Ito, M., «Depresia cerebeloasă pe termen lung: caracterizare, transducția semnalului și roluri funcționale», în *Physiology Review,* nr. 81, 2001, pag. 1.143-1.195.

—, «Organizarea moleculară a depresiei cerebeloase pe termen lung», în *Nature Reviews Neuroscience,* nr. 3, 2002, pag. 896-902.

Ito, M. și M. Kano, „Depresia de lungă durată a fibrelor paralele: transmiterea celulelor Purkinje indusă de stimularea conjunctivă a fibrelor para Ilei și a fibrelor cățărătoare în cortexul cerebelos", în *Neuroscience Letters,* nr. 33, 1982, pag. 253-258.

Koch, C., *În căutarea conștiinței,* Roberts, CO, 2004.

Lalonde, R., AN Bensoula y N. Filali, «Rotorod sensorimotor lear ning in cerebelar mutant mice», en *Neuroscience Research,* nr. 22, 1995, pag. 423-426.

McCormick, DA și RF Thompson, „Cerebelul: Implicare esențială în răspunsul pleoapei condiționat clasic", în *Science,* nr. 223, 1984, pag. 296-299.

Marr, D., «O teorie a cortexului cerebelos», în *Journal of Physiology,* nr. 202,

1969, pag. 437-470.

—, «Memoria simplă: o teorie a arhicortexului», în *Philosophical Transactions of the Royal Society of London B Biological Sciences,* nr. 262, 1971, pag. 23-81.

Masland, RH, „Planul fundamental al retinei", în *Nature Neuroscience ,* nr. 4, 2001, pp. 877-886.

Polyak, S.L., *Retina,* Chicago, University of Chicago Press, 1941.

Ramon y Cajal, S., Histology *of the Nervous System* (1904), Oxford, Oxford - University Press, 1995 (ed. spaniolă: Histología *del sistema nervio del hombre y de los vertebrates,* Madrid, Consejo Superior de Investigaciones Científicas, 1911).

Sieburth, D., O. Ch'ng, M. Dybbs, M. Tavazoie, S. Kennedy, D. Wang, D. Dupuy, J. E Ruai, DE Hill, M. Vidal, G. Ruvkun y JM Kapjan, «Systematic analysis of genes required for synapse structure and function», en *Nature,* n° 436, págs. 2005. 510-517.

Smirnakis, SM, AA Brewer, MC Schmid, AS Tolias, A. Schuz, M. Augath, W. Inhoffen, BA Wandell și NK Logothetis, «Lipsa reorganizării corticale pe termen lung după leziunile retiniene la macaci», în *Nature,* nr. 435, 2005, pag. 300-307.

Vaney, DI și WR Taylor, «Selectivitatea direcției în retină», en Opinia *actuală în neurobiologie,* nr. 12, 2002, pag. 405-410.

Welsh, JP, H Yamaguchi, XH Zeng, M Kojo, Y Nakada, A Takagi, M Sugimori și RR Llinas, „Învățarea normală în timpul prevenției farmacologice a depresiei pe termen lung a celulelor Purkinje", în *Pro ceedings of the National Academy of Sciences, SUA,* nr. 17.166-17.171.

Young, JZ, *Programele creierului,* Oxford, Oxford University Press, 1978.

Zeki, S., „Splendori și mizerii ale creierului", în Tranzacții *filozofice ale Societății Regale din Londra, Științe Biologice,* nr. 354, 1999, pp. 2.053-2.0

CĂUTAREA CONTINUĂ

Știința și filosofia în căutarea Prințului Rațiune

DANIEL ROBINSON

În cercurile intelectuale din Marea Britanie de la începutul secolului al XVII-lea, exista un mare interes pentru anatomie. Marele William Harvey , care se întorsese din Italia în 1602, avea să țină prelegerile sale de pionierat , Lumley Lectures, începând cu 1615. Cambridge a fost unul dintre centrele acestui interes reînnoit pentru funcționarea corpului. Harvey părăsise Cambridge în 1600, cu diploma în mână, pentru a primi instruire la Padova chiar de la Fabricius, în timp ce tânărul său coleg Phineas Fletcher (1582-1650) își încheia propria carieră academică la King's College. Știm cu toții cine a fost Harvey, dar Fletcher zace aproape uitat în negura trecutului. Niciun student la anatomie nu ar schimba metoda de instruire a lui Fletcher cu cea a lui Harvey. Cu toate acestea, curiozitatea care ne determină să reînviem abordarea lui Fletcher este răsplătită din plin, deoarece modul în care înțelegem lumea naturală, sau o parte a ei, este strâns legat de metodele selectate pentru această sarcină.

Publicată în 1633, * *Insula Purpurie** de Phineas Fletcher este o alegorie în douăsprezece cântece care îl ghidează pe cititor prin teritoriul misterios al corpului uman. Insula își datorează culoarea materiei purpurii cu care Dumnezeu a creat noul pământ. Cântecul

Descoperirea anatomică este cântată de Thirsil unui public de tineri păstori. Dacă ne imaginăm că acești băieți seamănă cu studenții de astăzi, vom fi de acord că interesul lor nu este stârnit decât atunci când Thirsil ajunge la cântul 6, căci în strofa 28 intrăm în domeniul...

> Prințul Insulelor, cu o constituție mai mult decât cerească,
> [...] numit pe bună dreptate *Intelectul* atotvăzător ;
> Toate de o natură strălucitoare și glorioasă, căci nimic nu este pământesc;
> A cărui față seamănă cu soarele și aspectul cel mai divin
> Nicio viziune umană nu poate descrie vreodată: Căci atunci când cineva îl privește,
> Este copleșit de stânjeneală sau surpriză față de o asemenea măreție.

Apoi, ajungând la strofa 30, nu există nicio îndoială cu privire la compoziția Prințului:

> Trupul său ciudat nu este corporal,
> Dar materie fără materie; niciodată plină,
> Nici măcar atunci când este plin; deși are la îndemână
> Toate cerurile și tot pământul și tot ce este în ele;
> Dar ar putea conține mii și mii de ceruri,
> Și să rămână la fel de gol ca la început;
> Și cu cât ia mai mult, cu atât este mai dispus să ia mai mult. [1]

Harvey și Fletcher se angajaseră în străvechea căutare a acelui sediu al sufletului rațional, a acelui „loc al formelor" pe care Aristotel, cu înțelepciune, nu a încercat să-l localizeze, a acelui „corp cel mai straniu" ale cărui proprietăți definitorii par a fi orice altceva decât corporale . După cum arată clar acest volum, căutarea continuă și, după cum arată și mai clar, există mai puțină certitudine astăzi decât în trecut cu privire la ce este probabil să fie descoperit și unde este cel mai probabil să fie descoperit.

Printre coautorii acestui volum se numără un distins om de știință și un grup

de filozofi realizați și influenți. Pentru a situa diferențele care le animă schimburile în contextul mai larg al istoriei ideilor, este util să luăm în considerare încă o dată metodele lui Harvey și Fletcher, ambii pornind într-o călătorie a descoperirii, ambii dedicați unui mod de explicație aparent validat de practici umane cu valoare dovedită. Privite în această lumină și cu rezervele pe care le voi sublinia în mod corespunzător mai târziu, aș spune că John Searle și Daniel Dennett ar dori să fie asociați cu Harvey, chiar dacă filozofiile lor speculative fac parte de fapt din tradiția Fletcher. La fel ca Fletcher, ei sunt studenți ai anatomiei timpului lor, folosind-o pentru a spune o poveste. Depinde de cititor să interpreteze concluziile lor ca alegorii sau ca titluri de ziar. Cu toate acestea, ca narațiuni, ele nu ar trebui confundate cu misiunea foarte diferită a științei , atât experimentală, cât și teoretică.

Max Bennett și Peter Hacker, atrăgând cu abilitate atenția asupra acestui aspect, ar insista cu tenacitate că face parte din însăși natura cazului că acestea sunt căi diferite care duc la obiective respectabile, dar fundamental diferite. Lăsând deoparte orice judecată asupra valorii relative a diferitelor moduri de cercetare și explicație, am putea sublinia că paginile lui Bennett și Hacker (cu excepția pasajelor explicit tehnice ale lui Bennett) sunt înrădăcinate în tradițiile de mult acceptate ale filosofiei analitice. Prin aceasta nu mă refer la o presupusă „descoperire" a unei mâini de filozofi simpliști de la Oxford, ci mai degrabă la acea misiune centrală a dialogurilor lui Platon - clarificarea termenilor, formularea problemelor sub formă argumentativă, exigența de coerență și consecvență. La toate acestea, Aristotel ar adăuga conținutul lumii naturale și lărgirea misiunii filosofice care rezultă din această adăugare. Dennett și Searle au povești de spus, și sunt povești bune spuse de maeștri în domeniu. Bennett și Hacker, în importanta lor carte, *Fundamentele filosofice ale neuroștiinței,* Ei au concluzionat că valoarea de adevăr a unor astfel de povești nu putea fi evaluată deoarece alegerea termenilor era particulară , neștiințifică și străină filosofiei. Ceea ce au găsit la Searle , Dennett și mulți alți lideri actuali în științele cognitive a fost pur și simplu un exces față de ceea ce Gilbert Ryle a făcut aluzie ironic prin sintagma „ea a venit într-o mare de lacrimi și într-un palanchin". †††În această carte, cititorul este trimis înapoi la acele

†††O aluzie la gustul pentru ornamentație retorică, dramatism și superficialitate

povești într- o formă condensată și instructiv dialectică.

Reflectând în propriul capitol, Bennett povestește o întâlnire cu John Eccles, care a insistat sardonic că cercetarea în neuroștiințe necesită întotdeauna „un sudor bun". Bennett își exprimă apoi propria convingere că „fiecare laborator de neuroștiințe cognitive de primă clasă are nevoie astăzi de un filosof critic și analitic bun". După mulți ani de muncă în laborator și reflecție, știu din proprie experiență că Eccles avea dreptate, dar nu sunt atât de sigur că Bennett are. Îndoiala mea provine dintr-un scepticism mai general cu privire la oportunitatea combinării unor discipline bine definite, fie ele legate între ele sau nu. Odată ce o disciplină bine definită, precum etica, se recombină în ceva numit bioetică, există tendința de a crede că trebuie să căutăm un precept etic mai profund care să acopere cazul furtului de ficat, deoarece preceptele care acoperă furtul de mașini sunt considerate insuficiente pentru această sarcină. La rândul său, în compusul „ neuroștiință cognitivă " , adjectivul în sine pune o problemă rezistentă analizei filosofice, cel puțin din zilele în care Platon i-a dat glas lui Socrate din *Urlătorul. Prima întrebare* pe care și-o pune Bennett — cum să derivăm psihologicul din sinaptic — este și întrebarea lui Fletcher. Ceea ce i-a lipsit lui Fletcher în încercarea sa de a *-l vedea* pe Prinț, Harvey nu a putut oferi. Nici bisturiul, nici laparoscopul nu sunt de vreun folos aici. Există argumente solide pentru a crede că misiunea extinsă a *psihologiei cognitive...* O abordare realistă și informată trebuie să urmeze o cale diferită de cea care ghidează progresul omului de știință ca pelerin.

Poate că îmi pot contura mai bine ideea referindu-mă chiar la cercetarea pe care o rezumă Bennett , care concluzionează că rezultatele percepției vizuale sunt mai corelate cu procesele corticale „superioare" decât cu cele mai apropiate de originea lor din retină. Există un corpus lung și consistent de descoperiri care stabilesc că, cel puțin în auz și vedere, „acordarea" sistemului devine din ce în ce mai precisă pe măsură ce evenimentele se deplasează de la nivelul neuronilor primari la destinațiile lor corticale finale. Dar fenomenul rivalității binoculare este diferit de acordarea sistemului la benzi de frecvență mai înguste. Există o

din partea ortodoxiei intelectualiste în filosofia contemporană a minții . *(Nota traducătorului)*

proprietate fenomenologică evidentă conform căreia comportamentul condiționat al animalului semnalează prima dată când o imagine și apoi cealaltă alternează în dominanța perceptivă. Întrebarea care apare în mod natural în acest cadru se referă la modul în care toate acestea au loc la creaturile cu ochi poziționați lateral, astfel încât același tip de *rivalitate binoculară* nu poate exista . Animalele cu ochi poziționați medial se confruntă cu un spațiu vizual în care același obiect poate concura pentru recunoaștere . Cei cu ochi poziționați lateral se confruntă cu două câmpuri vizuale separate care nu au niciun obiect în comun.

De ce menționez asta? Fac asta pentru a sublinia ideea evidentă (deși de obicei trecută cu vederea) că nu numai că există un *animal real* care vede ceva, dar modul în care mediul vizibil îi influențează percepția depinde de o realitate ecologică mult mai largă și mai complexă, la care creaturile de un anumit fel trebuie să se adapteze. Acest fapt impune limite de diferite grade de severitate generalizărilor între specii și nu pare nerezonabil să suspectăm că, atunci când aceste generalizări includ conștiința „tranzitivă" sau „intranzitivă", intră în joc limite și mai stricte. Așa cum peștele nu poate descoperi apa, observațiile făcute în bula artificială care este un laborator sunt iremediabil îndepărtate de viața reală trăită în lumea vizibilă. Cu alte cuvinte, este mult mai ușor să determini cum e să fii pisică decât să -ți dai seama cum e să nu fi văzut nimic altceva decât proiecții menite să excite celulele retiniene, pe parcursul unei întregi vieți petrecute într-o cușcă de laborator.

Printre îndoielile exprimate de Bennett și Hacker față de neurocercetătorii cognitivi care altfel erau deasupra oricărei suspiciuni, cea care iese în evidență este în special cea referitoare la afirmația că nu suntem întotdeauna conștienți de ceea ce percepem. Bennett observă că „cineva poate percepe un obiect X fără să fie conștient de el, fie pentru că îl identifică în mod eronat ca fiind Y, fie pentru că atenția sa nu este atrasă în mod constant către el, ceea ce s-ar putea datora faptului că nici măcar nu îl percepe sau pentru că se *concentrează intenționat asupra lui".* Acest lucru nu este foarte convingător. Identificarea greșită a unui obiect nu implică deloc că ceva nu a fost perceput conștient. „Identitatea" oricărui obiect vizual nu este stabilită în mod univoc. Dragostea pe care Oedip a simțit-o pentru Jocasta nu a fost filială. Floarea albastră, care strălucește mai puternic în lumina lunii decât cea galbenă, care la rândul ei strălucește mai puternic la amiază, nu este

o dovadă împotriva conștiinței perceptive tranzitive. Nici mărturia albinei, a cărei sensibilitate spectrală maximă se află în banda ultravioletă a spectrului electromagnetic. În plus, un indicator fiabil al funcționării conștiinței perceptive tranzitive este segregarea *intenționată a* elementelor în spațiul vizual. Cât despre perceperea lui X „fără a fi conștient de acest lucru", mă tem că un consens general în această privință ar necesita un fel de apărare teoretică specială.

Acestea sunt scrupule destul de minore, mai ales în comparație cu cele ridicate de John Searle. Scopul său este de a plasa din nou conștiința în creier, acolo unde Galen și filozofii hipocratici au localizat-o cu mult timp în urmă. Searle îi depășește în concluziile sale, beneficiind de progresul științific și prin distincția dintre *stare* și *loc.* Se presupune că afirmația că conștiința este o *stare* a creierului este mai puțin particulară decât afirmația că aceasta se află „în" creier.

Utilizarea stărilor și a termenilor înrudiți, precum procese și mecanisme, a devenit atât de frecventă încât acești termeni se bucură acum de un statut aproape protejat. Dar nu este complet clar dacă servesc vreunui alt scop decât acela de a impune ilicit concluzii asupra unui argument care nu a fost încă formulat corespunzător. Și aceasta este o practică obișnuită. Searle își permite să postuleze „stări" și adaugă imediat ceva numit „ caracterul calitativ al consumului de bere", care, spune el, „este diferit de caracterul calitativ al ascultării *Simfoniei a IX-a* de Beethoven ". Sunt destul de sigur că știu diferența dintre a bea bere și a asculta Beethoven, dar sunt mult mai puțin sigur că oricare dintre aceste activități are un „caracter calitativ". Nu am consumat niciodată un caracter calitativ, deși am băut destul de multă bere. Este aceasta o obiecție minoră? Poate. Dar în acest domeniu, obiecțiile minore sunt cumulative și pot atinge echivalentul filosofic al unei mase critice.

Searle ajunge la esența criticii sale cu concizia sa admirabilă, caracteristică:

> O mare parte din cele mai bune lucrări din neuroștiințe reprezintă un efort de a explica modul în care procesele cerebrale cauzează experiența vizuală și unde și cum are loc această experiență în creier. În mod uimitor, Bennett și Hacker neagă existența experienței vizuale în acest sens, în sensul quale .

Înainte de a lua în considerare negarea uimitoare a lui Bennett și Hacker , este important să examinăm ceea ce consider o afirmație mult mai uimitoare: aceea că cea mai bună lucrare din neuroștiințe promite să explice modul în care procesele vizuale *cauzează* experiența vizuală (care are loc undeva în creier). John Searle este perfect conștient că întreaga problemă a cauzalității este esențială pentru problema în cauză. Prin urmare, el știe că afirmația că o astfel de cauzalitate este bine stabilită - bazată pe studii solide care demonstrează cum funcționează totul - nu rezistă examinării filosofice. Nu există un acord unanim cu privire la termenii relației cauzale (sunt fapte, obiecte materiale, termeni conceptuali, evenimente, condiții ?) sau chiar dacă aceștia trebuie să existe invariabil. La urma urmei, motivul pentru care Joan a supraviețuit a fost că nu a băut otrava. Aici, „cauza" supraviețuirii este un non-eveniment . Ca să ajungem la subiect, dacă tot ceea ce definește domeniul...

 Dacă

starea mentală 223 este într-adevăr rezultatul cauzal al unor seturi de „ stări” - cerebrale , atunci, așa cum spune maxima, *fizica ca știință este completă* , iar filosofii ar trebui să ia în considerare recalificarea într-o altă carieră.

Ca locuitor de lungă durată al planetei Pământ, nu am nicio îndoială că o organizare sănătoasă și funcțională a corpului , mai ales dacă este inclus și sistemul nervos, constituie *condiția necesară* pentru ceea ce ne place să numim viața noastră mentală, cel puțin în încarnările sale sublunare. Cu toate acestea, sugestia că țesutul nervos *provoacă* toate acestea ar provoca, cel puțin, uimire într-o epocă în care știința nu degenerase încă într-un fel de retorică. Așa cum gânditorii de marcă din domeniul fizicii sunt disciplinați și precauți în ceea ce privește cauzalitatea, apar neurocercetătorii cognitivi și lacheii lor filosofici, întrebându-se cum ar putea cineva ezita să adopte un obiectiv atât de evident: „cum procesele cerebrale provoacă experiența vizuală și unde și cum se întâmplă aceasta *în* creier”. Chiar presupunând că am putea susține, într-un sens metafizic acceptabil , că am stabilit cum forțele gravitaționale *determină* cheia ușii să cadă spre centrul Pământului, explicația funcționează (dacă funcționează deloc) doar în măsura în care cheile Pământului și cele ale casei au ambele masă, iar magnitudinea separării lor poate fi specificată în mile, picioare, inci sau (cu reticență) metri. Totuși, ștacheta metafizică este ridicată considerabil mai sus atunci când legătura cauzală este cea dintre activitatea metabolică *de undeva* , pe de o parte, și ascultarea *liedului* „An die Freude” - ca să nu mai vorbim de compunerea lui - pe de altă parte.

Searle își exprimă îndoieli serioase cu privire la amploarea jocului lingvistic . Resursele conceptuale pe care le utilizează Bennett și Hacker...

Sunt în mod clar preluate din Wittgenstein, iar Searle ridică ceea ce este, de fapt, o rezervă binecunoscută cu privire la, în cuvintele sale, eroarea wittgensteiniană de a confunda

> criteriile de aplicare a conceptelor mentale cu stările mentale în sine. Cu alte cuvinte, ei confundă criteriile comportamentale pentru *atribuirea* predicatelor psihologice cu *faptele atribuite de aceste* predicate psihologice, ceea ce reprezintă o eroare gravă.

Problema este prea abstractă pentru a fi abordată sumar . Nu se poate contesta faptul că relatările la persoana întâi și la persoana a treia despre durere provin din surse diferite. Este incontestabil că baza pe care Pérez *simte* o durere de dinți este diferită de baza pe care Martínez consideră că Pérez are dureri. Cu toate acestea, ceea ce este controversat este afirmația că simpla prezență a unor modele de descărcare excesivă în fibrele relevante ale ramurii maxilare a nervului trigemen justifică afirmația că Pérez simte „durere" exact în același sens în care ar simți-o persoanele competente social și lingvistic din cultura sa . Am îndoieli în această privință, dar, din nou, lipsa de spațiu mă împiedică să fac o expunere mai completă. Cu toate acestea, Searle , după ce își observă datoria față de Wittgenstein, preia un pasaj din Bennett și Hacker doar pentru a-l interpreta într-un mod ciudat. Acesta este pasajul:

> Criteriile pentru atribuirea unui predicat psihologic sunt [...] parțial constitutive ale *semnificației* acelui predicat [...] Creierul nu îndeplinește criteriile pentru a fi un posibil subiect al predicatelor psihologice.
>
> (p. 83)

Searle interpretează aceste cuvinte ca o negare a conștiinței creierului, bazată pe faptul că acesta nu se poate „comporta " în niciun fel. El spune că teza principală a lui Bennett și Hacker este că „creierul este incapabil să manifeste un comportament adecvat". Într-adevăr, asta spun Bennett și Hacker. Cu toate acestea, Searle trece cu vederea argumentul subtil (poate prea subtil) din care trag această

concluzie. Nu este vorba că conștiința nu poate fi atribuită creierului pentru că acesta este incapabil să manifeste un comportament adecvat; mai degrabă, atribuțiile în cauză, pentru a fi semnificative în sensul strict al cuvântului, trebuie să îndeplinească aceleași criterii ca orice predicat. Afirmațiile că Pérez este înalt, creierul este ud și Enriqueta este tânără sunt inteligibile doar în măsura în care „ud", „tânără" și „înaltă" nu sunt scoase din cutia etichetată GÂNDACĂ, care, în plus, este vizibilă doar persoanei care o ține. Sensul în care Pérez, ca individ izolat, ar putea să nu atribuie sens adjectivului „înalt" pe care alții îl folosesc pentru a-l descrie este același sens în care ar putea spune în mod impropriu că ceva „doare", *chiar și atunci când vorbește despre sine.* Dacă nu cumva i-am interpretat și eu greșit pe Bennett, Hacker și Wittgenstein, concluzia nu este că creierul nu poate fi conștient , ci mai degrabă că afirmația că este conștient este la fel de incomprehensibilă ca afirmația că este social-democrat.

În diverse momente, dar mai ales spre sfârșitul interesantului său eseu, Searle face aluzie la posibilele contribuții ale științei experimentale la problemele filosofice. El observă că nu este previzibil ca întrebările legate de „viața bună" etc. să beneficieze de un astfel de beneficiu; cu toate acestea, se așteaptă ca unele probleme filosofice să cedeze într-adevăr descoperirilor științifice. Este păcat că exemplul specific pe care îl oferă ar putea lăsa mai mult de un cititor iremediabil nedumerit. Acestea sunt cuvintele lui Searle :

> Nu fac o distincție clară între întrebările științifice și cele filozofice . Permiteți-mi să dau un exemplu pentru a explica cum pot descoperirile științifice să-mi ajute munca filosofică . Când ridic brațul, intenția mea conștientă în acțiune provoacă o mișcare fizică a corpului meu. Dar mișcarea are și un alt nivel de descriere, deoarece este cauzată de o secvență de descărcări neuronale și de secreția de acetilcolină la nivelul plăcilor terminale ale axonilor neuronilor motori . Pornind de la aceste fapte, pot face o analiză filosofică pentru a arăta că unul și același eveniment trebuie să fie atât un eveniment conștient calitativ, cât și subiectiv și să posede, de asemenea, multe proprietăți chimice și electrice . Dar aici se termină analiza filosofică. Acum trebuie să știu mecanismul exact care face ca acest lucru să funcționeze.

Citind aceste rânduri, ne putem întreba ce fel de „analiză filozofică ” duce la concluzia că mișcarea brațului are „multe proprietăți chimice și electrice”. La un nivel, este evident (și ar fi fost evident și pentru locuitorii peșterilor analfabeți) că brațele au greutate, că ceva sub pielea brațului se încordează și că aceleași brațe, atunci când sunt străpunse de obiecte ascuțite, emană un lichid roșu, fierbinte. Nu există nicio îndoială că este sarcina științei experimentale să stabilească detaliile tuturor evenimentelor subcutanate asociate cu ridicarea unui braț. Este îndoielnic că o analiză filozofică ar fi ceva mai mult decât o distragere a atenției în timp ce echipa de cercetare este angajată în această muncă importantă. Între timp, cei cu înclinații filozofice s-ar putea întreba pur și simplu care este diferența dintre cineva căruia i se ridică brațul și obținerea aceluiași rezultat ridicându-și intenționat brațul . Este plauzibil să concluzionăm, fără nicio investigație științifică, că o astfel de diferență va fi exprimată undeva sub piele, recunoscând totodată că diferențele psihochimice nu vor „explica” intențiile. Însă când acceptăm judecata prudentă, bine întemeiată și informativă conform căreia filosofii s-au dedicat chestiunii abstracte a activității volitive - cu mult înainte ca cineva să știe despre neuroni sau potențiale ale plăcilor terminale - afirmația că acești autori trebuiau să cunoască „mecanismul exact care face ca aceasta să funcționeze” pare greu de crezut.

Searle este dispus să accepte concluziile lui Wittgenstein cu privire la problema „jocului lingvistic”, dar le consideră irelevante pentru proiectul neuroștiinței cognitive . Astfel,

> Când investigăm ontologia durerii — nu condițiile pentru a juca jocul limbajului, ci ontologia fenomenului în sine — putem uita de comportamentul extern și ne putem concentra pur și simplu pe a descoperi cum creierul provoacă senzații interne.

Fără a uita îndoielile menționate anterior cu privire la „cum provoacă creierul [...] senzații”, să luăm în considerare sensul în care Searle vorbește despre „ontologia durerii”, referindu-se la senzația „reală” în sine. Un creier viu nu este niciodată inactiv și, în consecință, numărul de corelații neuro-fenomenale este practic nelimitat. Așa-numitele căi clasice ale durerii se termină în talamus, fără a

exista un „centru” cortical al durerii ca atare. Prin urmare, întrebarea s-ar reduce la modul în care nucleii talamici „provoacă” durere. Nucleii sunt constelații de corpuri celulare integrate care funcționează ca o unitate. Astfel, putem rafina și mai mult căutarea: cum provoacă durere potențialele de acțiune postsinaptice emise de corpurile celulare din talamus?

Să presupunem, de exemplu, că întindem sau răsucim brațul cuiva până la punctul în care simte durere. Durerea, desigur, este simțită în braț, nu în talamus, deoarece nu există nimic în creier care să „simtă” ceva. Știm că Pepe suferă pentru că tresare și spune: „Au!”. De asemenea, știm că fibrele C sunt activate și că semnalele care intră pe suprafața dorsală a măduvei spinării vor călători către creier și nucleii talamici relevanți. Dar multe alte lucruri, provenind din același braț, fac aceeași călătorie. În plus, „au!”-ul nu apare până când nivelurile de descărcare nu ating și nu depășesc o valoare critică. Fără o reacție facială , posturală sau vocală din partea lui Pepe, am avea toate aceste date neurofiziologice, dar nu am ști ce să facem cu ele. În cele din urmă, Pepe este cel care are ultimul cuvânt despre durerea sa . Dar cum știm care dintre „semnalele” lui Pepe este cel care indică durerea? Și cum știe Pepe asta? Cred că Searle a abandonat jocul lingvistic cu cel puțin un fonem prea devreme.

În discuția sa despre eroarea mereologică, Searle respinge existența oricărei erori și susține că nici măcar argumentele wittgensteiniene nu pot evita noțiunea că (într-un sens special) „creierul gândește” sau „creierul vede”. În cuvintele lui Searle, Bennett și Hacker susțin că creierul nu gândește și că gândirea nu poate avea loc în creier. El insistă că „pentru a demonstra că creierul nu poate fi locul unor astfel de procese ar fi nevoie de un argument independent, un argument pe care nu l-am putut găsi”. Mai degrabă, este...

> Tot ce impune argumentul wittgensteinian este ca creierul să facă parte dintr-un mecanism cauzal în cadrul unui sistem global capabil să producă - comportament. Și această condiție poate fi îndeplinită chiar și atunci când anumite procese psihologice sunt localizate în creier.

Pe măsură ce obiecțiile mărunte se acumulează, începem să percepem că a

vorbi despre gândire ca despre un „proces” necesită, mai mult sau mai puțin, găsirea locului „său” și că doar un contraargument independent va slăbi ceea ce dictează bunul simț în această privință. Gândirea ca „proces” ar fi, probabil, desigur, un „proces cerebral”, cel puțin dacă trebuie să alegem dintre diversele organe ale corpului. Totuși, dacă „gândirea” este cuvântul pe care îl aplicăm acelei cantități vaste de evenimente, așteptări, credințe, judecăți, strategii și așa mai departe, comprimate într-un minut neobosit de conștiință - sau dacă o aplicăm uneia dintre acele idei cu care suntem obsedați neîntrerupt perioade lungi de timp - aș spune că sarcina probei revine celor care susțin că oricare dintre aceste lucruri este un „proces” și cu atât mai mult că are loc în creier. Să luăm în considerare *informația,* acea entitate fără masă care nu ocupă niciun spațiu, în sensul strict al ceea ce modifică probabilitățile și entropia generală a unui sistem. Nu este momentul să intrăm în lumea fascinantă a indeterminației și suprapunerii cuantice și să încercăm să salvăm pisica lui Schrödinger, dar merită totuși să recunoaștem că știința noastră cea mai dezvoltată este mult mai puțin dedicată ideii că efectele reale necesită locuri, mase și „procese ” observabile. S-ar putea spune că statutul definitiv al fizicalismului depinde tocmai de cât de bună și mai cuprinzătoare este explicația vieții mentale, dar cu siguranță este prea devreme pentru a lua o poziție fermă în această privință. Punctul de plecare corect este să începem cu propria noastră alegere de termeni, asigurându-ne că nu adoptăm moduri de vorbire care exclud practic posibilitatea de a alunga ignoranța noastră sistematică.

Daniel Dennett, cu care John Searle a avut propriile sale dezacorduri, este totuși la fel de critic față de Bennett și Hacker ca și Searle, dar din motive diferite, dacă nu chiar complet diferite . Principala linie de apărare a lui Dennett este să se prezinte ca cineva care extinde de fapt gândirea „Sfântului Ludwig” (așa cum îl numește Dennett), aplicând-o comportamentului roboților, computerelor care joacă șah și chiar creierului și părților acestora - comportamente suficient de similare cu cele ale oamenilor pentru a permite aplicarea termenilor psihologici. Citând propriile sale scrieri timpurii, Dennett susține că tocmai pentru că există două „niveluri de explicație”, suntem chemați la sarcina de a le corela, o sarcină care necesită analiză filosofică. *{Conținut și conștiință}* (pp. 95-96). Totuși, faptul că există într-adevăr două niveluri de explicație nu garantează în sine că acestea sunt

sau pot fi legate, nici, dacă sunt, că această relație este cauzală. Dacă relația care trebuie examinată este cea dintre, de exemplu, temperatura ambiantă și energia cinetică medie a sistemului, am avea o relație de identitate. Dar dacă relația este cea dintre un număr de stradă și reședința unei anumite familii, este clar că faptul că familia Pérez locuiește la numărul 77 de pe Paseo de los Robles, iar familia Martínez la numărul 79, nu are nicio valoare informativă pentru scopurile noastre. Există o relație clară între decizia lui Guillermo de a participa la concert și direcția în care se mișcă picioarele sale după ce parchează mașina. Există o relație la fel de clară între anumite activități din canalele extrapiramidale ale guvernatorului Morris, pe de o parte, și acțiunea sa de a tăia „Noi, Statele" și de a scrie în locul acestuia „Noi, Poporul" în preambulul Constituției Statelor Unite, pe de altă parte. Totuși, nu ar fi grotesc să sugerăm că, atunci când se ia în considerare individualizarea drepturilor de care se bucurau primii cetățeni ai Statelor Unite, trebuie luate în considerare două niveluri de explicație, iar unul dintre ele se referă la canalele extrapiramidale ale guvernatorului Morris? Mi se pare absurd.

Și cum rămâne cu Kasparov și computerul care se presupune că „joacă șah"? A venit un moment în care adversarul frustrat al lui Deep Blue a declarat că adversarul său pur și simplu nu *juca* șah. Îi lipsea pasiunea, era imun la presiune și nu se implica direct în el. Reamintiți-vă de *„Scrisorile despre educația estetică a omului" de Schiller* , unde ni se spune că omul nu este niciodată atât de autentic el însuși ca atunci când joacă. Luați în considerare factorii culturali și dispozițiile comportamentale largi și diverse care trebuie recrutați pentru a califica o anumită activitate drept „joc" și apoi comparați-i cu orice „proces" care îl face pe Deep Blue să mute nebunul la b2. Deep Blue „joacă" șah doar în același sens în care un cuptor cu microunde „gătește " supă, chiar dacă programarea este mult mai complexă.

Este posibil să fim tentați să spunem că, dacă aceasta este caracterizarea corectă, atunci Kasparov joacă și șah în același sens în care un cuptor cu microunde gătește supă, chiar dacă programarea este mult mai complexă? În cele din urmă, acesta este motivul pentru care teza IA puternică este atât de interesantă . Dacă adoptăm „atitudinea intențională" a lui Dennett, este nu numai permis, ci și avantajos din punct de vedere conceptual, să recunoaștem că Deep Blue posedă orice motive, sentimente, credințe și atitudini pe care le considerăm potrivite pentru

a explica ceea ce face Kasparov. În acest fel, Kasparov nu este „redus” la nivelul unei mașini, ci mai degrabă mașina este ridicată la rangul de sisteme inteligente. Dacă atât Kasparov, cât și Deep Blue sunt ascunși în spatele unui ecran, iar întrebările relevante ale testului Turing primesc aceleași răspunsuri de la amândoi, atunci amândoi sunt „inteligenți” în ceea ce privește nivelul întrebărilor la care se răspunde corect. Dar apoi intrăm în Camera Chineză a lui Searle și începem să credem că Deep Blue nu este nimic mai mult decât un dispozitiv pentru sortarea plăcuțelor și că „răspunsurile” sale nu sunt răspunsuri propriu-zise, ci pur și simplu „ieșiri ”. Dezbaterea se prelungește și mai mult, dar numai din cauza unei forme de isterie intelectuală care îi face pe oamenii extrem de educați incapabili să recunoască prostia propoziției inițiale, și anume că Deep Blue/Wgzz *se pricepe la șah.*

Dennett atrage atenția asupra preferinței lui Hacker pentru „sens” și „nonsens” ca trăsături relevante ale unui argument filosofic, în contrast cu „adevărul” sau „falsul” științei. Clasificările lui Hacker , poate excesiv de sobre, sunt costume prea strâmte pentru purtarea de zi cu zi. Oricare ar fi produsul din atelierele neobosite ale științei, însuși faptul că este supus unor corecții, revizuiri și ajustări ulterioare arată clar că, pentru început, nu era „adevăr ”; nici nonsens, sau cel puțin un nonsens complet , cu excepția câtorva cazuri evidente. (Obiectele încălzite nu se ridică pentru că dobândesc substanța *ușurinței,* iar ochii uscați nu sunt un semn de vrăjitorie.) Prostiile, în plus, reprezintă o condamnare prea severă pentru pozițiile filosofice obtuze, prea ambițioase, gratuit autobiografice sau prea serioase . (În perioada facultății, consideram că încercarea lui Hume de a reduce conceptul de cauzalitate la coincidența constantă a anumitor obiecte în experiență era o glumă scoțiană tipică. Abia mai târziu am fost nevoit să accept concluzia dezamăgitoare că vorbea serios!) Dar o critică adecvată a clasificărilor lui Hacker nu se poate baza doar pe iritarea provocată de alegerea cuvintelor sale. Înțelepții Cercului de
La Viena, membrii zeloși ai Societății Ernst Mach erau înclinați să considere orice afirmație non-empirică drept un nonsens literal . Ingenioșii membri ai Clubului Scriblerus - Pope, Swift, Arbuthnot - după ce au citit paginile lui Locke despre identitatea personală, au ajuns la concluzia că acesta își pierduse cumpătul filosofic . Să fim de acord că tratatele filosofice pe teme importante trebuie să se străduiască

să fie clare, accesibile, coerente și, din păcate, respectuoase față de subiectele alese. Cei care eșuează într-o astfel de încercare jignesc bunul simț. Nu că bunul simț este arbitrul suprem; ci doar că arbitrul este cel care trebuie câștigat dacă tratatul vrea să exercite vreo influență dincolo de cei patru pereți ai seminarului. Dacă Spassky și Kasparov se îndoiesc că un computer poate „joca" șah, nu ar trebui Dennett să fie cel care să reconsidere problema?

În aceeași ordine de idei, Dennett pretinde că îi demască pe Ryle și Wittgenstein și demonstrează că (la fel ca Hacker) sunt nesinceri atunci când sugerează că există „reguli" care guvernează discursul filosofic , darămite discursul cotidian. El numește *eroarea categorică o cacealma.* Ryle și promisiunea sa privind o „logică a existenței". În ciuda întregii munci dedicate sintaxei, lingviștii încă își croiesc creierii pentru „Pisica a coborât din copac". Așa să fie. Nici nu se poate interzice „Micul dejun a fost o încântare pentru hipotalamus, având în vedere cât de sătul era conducta sa electrică pe măsură ce masa avansa ". Ceea ce este încălcat aici nu este o lege, ci o convenție - adică o *regulă* a cărei încălcare, spre deosebire de o lege, nu duce la o apariție în instanță, ci la o neînțelegere. Atunci când exprimările de acest fel devin obișnuite, neînțelegerile devin sistematice, încărcate de paradoxuri neintenționate și implicații neprevăzute, atenuate doar ocazional de o notă involuntară de umor.

Evident, „convențiile" Biroului Politic sunt diferite de cele ale Parlamentului britanic. Când vine vorba de stabilirea definitivă a sensului, întrebarea este exact ce convenții ar trebui să prevaleze. Aceasta ridică dilema de a decide câtă deferență să acorde psihologiei populare. Dennett îi avertizează pe neurocercetători să dea dovadă de „maxima prudență" atunci când manipulează termenii acestei psihologii, deoarece, așa cum spune el, „presupozițiile de utilizare le pot submina scopurile". Cum? Prin transformarea „ceea ce altfel ar fi teorii și modele empirice promițătoare în nonsensuri abia deghizate". Dacă înțeleg corect ce înseamnă aici „presupoziții de utilizare", aș spune că presupoziția centrală este că ceea ce Bennett și Hacker numesc „ descriere psihologică actuală" trebuie să facă posibilă întreaga gamă de interacțiuni practice și semnificative între vorbitori competenți. Este evident că, dacă argoul lor vulgar nu este nimic mai mult decât un nonsens abia deghizat , este îndoielnic că posedă resursele lingvistice pentru a beneficia măcar de iluminarea

filosofică. Poate că va veni ziua în care cetățeanul obișnuit va fi de acord să vorbească în termenii „ sistemului său cerebral triunic", abandonând cu umilință vechile, abia deghizate absurdități despre mere roșii, iarbă verde și cer albastru. Totuși, și în cazul în care omenirea s-a născut în acest joc lingvistic, ne întrebăm cum ar putea neurocercetătorii să împace vreodată ceva din „ sistemul cerebral triunic" cu – de fapt – lumea așa cum apare ea în realitate . Dacă discursul popular ar putea corupe cu adevărat teoriile și modelele empirice, ar face-o în primul rând pentru că teoriile și modelele nu au nicio relație specială cu chiar acest discurs, care este, în cele din urmă, discursul vieții reale. Poate că acest lucru ajută la explicarea motivului pentru care teoriile și modelele oferite astăzi nu sunt altceva decât modele de date, rezumate eficiente ale unor observații deschis antiseptice, atenuate de presupuse manipulări statistice și prezentate ca modele extrem de integrate ale... *nimănui,* nici măcar unui creier.

Pentru a se apăra împotriva acuzației de comitere a erorii mereologice, Dennett citează propriile sale scrieri anterioare în care stabilește o distincție între nivelurile de explicație personală și subpersonală. El vorbește despre sine ca despre un pionier în acest sens. Personal, i-aș fi acordat laurii lui Aristotel, care ne amintește că, pentru a explica furia, de exemplu, se poate vorbi despre modificări ale temperaturii sângelui sau, dimpotrivă, despre reacția la faptul că ai fost ignorat.

> Un fizician ar defini o stare a sufletului diferit față de un dialectician; acesta din urmă ar defini, de exemplu, furia ca dorința de a întoarce durere pentru durere [...], în timp ce primul ar defini-o ca o fierbere a sângelui sau a substanței fierbinți care înconjoară inima.[2]

De fapt, i-am putea atribui lui Aristotel primele rezerve cu privire la eroare, întrucât, în același tratat, el spune:

> A afirma că sufletul este cel care este mânios este la fel de inexact ca și cum ai spune că sufletul este cel care țese pânze de păianjen sau construiește case. Fără îndoială, este mai bine [...] să spunem că omul este cel care face toate acestea cu sufletul său. [3]

Distincția mai largă dintre explicațiile bazate pe cauze și cele bazate pe motive este, după cum se știe, venerabilă și nu mai puțin controversată pentru că este recunoscută pe scară largă ca fiind riguroasă din punct de vedere filosofic. Eroarea mereologică, însă, se manifestă în moduri diferite. Este transparentă atunci când se sugerează că tortul de ziua de naștere a fost tăiat de nervii eferenți de la nivelurile cervicale 5-8 ale măduvei spinării. Cu toate acestea, ea funcționează și atunci când motivul acțiunii lui Ronaldo este analizat ca suma unui număr mare de mini-motive. Astfel, explicația corectă pentru cumpărarea Toyota Prius de către Ronaldo ar fi că dorea un avantaj mecanic mai mare decât cel obținut prin mers pe jos. Faptul că Dennett este o victimă a acestei versiuni a erorii este evident cu o claritate lăudabilă în propriile sale cuvinte:

> Nu atribuim unor părți ale creierului o credință *completă* – aceasta ar fi o eroare. Nu. Atribuim acestor părți un tip de credință mai puțin important, atenuat.

Exemplul propus este cel al unui copil care „oarecum" crede că tatăl său este doctor. Acest lucru nu este deloc convingător. Pot exista unele îndoieli cu privire la obiectul credinței, dar cu siguranță nu există o „parte" a unei credințe. Cu toate acestea, exemplul în sine trebuie, în cele din urmă, să postuleze părți ale credințelor, deoarece „atenuarea" unei credințe prin acțiunea unui sistem fizic (de exemplu, părți ale creierului) înseamnă a-i schimba valoarea de-a lungul unui fel de continuum fizic, iar acest lucru, în cele din urmă, înseamnă a lucra asupra „părților" sale. Dacă a spune că creierul are credințe sună ca o proclamație a Reginei Roșii *din Alice în Țara Minunilor,* este și mai potrivit ca ea să afirme că o parte a creierului are credințe atenuate.

Dennett este deosebit de convingător atunci când respinge argumentele conform cărora creierul formează imagini de un fel sau altul. El afirmă pe bună dreptate că întrebarea dacă structurile interne ale creierului sunt sau nu organizate astfel încât să *funcționeze* ca producători de imagini este una empirică, dincolo de domeniul modurilor filosofice de analiză. Modul exact în care constelațiile interconectate de neuroni reacționează la lumea exterioară este o întrebare centrală

în neuroștiință și a fost abordată în mod deosebit de fructuos în cercetarea sistemului vizual. Nimeni nu crede de fapt că lumea vizuală este proiectată asupra creierului ca o *imagine* - și, bineînțeles, nu există imagini cu mirosuri sau sunete. Mai degrabă, cercetătorii caută relații izomorfe între caracteristicile optice ale lumii vizibile și modelele neuroelectrice asociate cu inspecția acesteia. Având în vedere toate acestea, ne confruntăm cu o întrebare foarte diferită: nu cum procesează sau „codifică" un anumit algoritm neuroelectric proprietățile optice ale lumii vizibile, ci relația dintre această codificare și ceea ce pretinde cel care percepe (vizual) că se întâmplă. A considera *această* întrebare potrivită pentru a ghida cercetarea empirică înseamnă a uita însăși esența întrebării, întrucât nimic din ceea ce se observă la nivelul neuroanatomiei funcționale nu „vede" în *vreun* sens , nici măcar într-un sens atenuat.

Este necesar mai puțin timp pentru a analiza apărarea lui LeDoux și a altor autori ai *nomenclaturii de către Dennett.* din neuroștiința cognitivă. A spune că „creierul" poate ști despre un pericol înaintea lui sau că noi știm ce este nu poate fi decât o corupere a limbajului și, la nivelul explicației științifice , o crestătură regretabilă în briciul lui Occam. Macacul nou-născut are celule în cortexul auditiv care reacționează la strigătele de suferință ale acestei specii. Aceste celule sau cortexul nu „știu" nimic, la fel cum un circuit rezistor-capacitate nu „știe" că a fost instalat un frigider, chiar dacă reacționează la o scădere de tensiune . În formarea creaturilor care se vor confrunta cu o lume periculoasă, există multă pre-cablare și o oarecare cablare permanentă . Natura le înzestrează cu capacitatea de a face din instinct sau reflexiv ceea ce nu poate aștepta până la finalizarea studiilor superioare. Acesta este echipamentul care ocolește *orice* învățare și, în consecință, orice cunoaștere. Dennett insistă asupra stocării în cache a oricărei cantități și varietăți de fapte într-un recipient conceptual prea elastic pentru a avea vreo formă și prea fragil pentru a suporta greutatea problemelor reale. Apărarea sa constă în evidențierea tuturor micilor fapte care pot fi înghesuite în acel recipient, dar, în general, simpla repetare a frazelor care au atras critici în primul rând este considerată ineficientă din punct de vedere retoric. El și mulți alții din mișcarea neuroștiinței cognitive *(* căci are toate caracteristicile unei mișcări) au adoptat un limbaj care este străin oricărei urechi în afară de a lor , dar își găsesc liniștea din

apariția aceloraşi fraze în toate cărțile și articolele lor. Justificarea lui Dennett pentru alegerea expresiilor deranjante, bazată pe frecvența cu care le folosește el însuși, este - pentru a adapta un exemplu din Wittgenstein - „ ca și cum cineva ar cumpăra mai multe exemplare ale ziarului de dimineață pentru a se asigura că ceea ce spune este adevărat”. [4] Și ce ziceți de Bennett și Hacker, în special de acesta din urmă? Am scris o recenzie foarte favorabilă a cărții sale în *Philosophy,* Și nimic din răspunsurile lui John Searle sau Daniel Dennett nu m-ar face să-mi reconsider judecata inițială. Am considerat acolo că scopurile autorilor erau tocmai cele care constituie adevărata misiune a filosofiei , care, în sensul cel mai larg, nu este nimic mai puțin decât critica vieții și, în ambițiile sale mai modeste, o anchetă critică asupra afirmațiilor noastre epistemice fundamentale. Istoria arată foarte clar consecințele schimbului acestei misiuni cu o poziție superioară în ierarhia în care se situează știința însăși. Există ceva foarte important în modul în care Sofocle, prin *Antigona,* apără fundamentele morale ale oricărei legi împotriva afirmațiilor regelui . Dar Sofocle nu a întreprins opera la care s-au dedicat Aristotel, Cicero, Toma d'Aquino și alți membri ai tradiției Dreptului Natural. Medicii hipocratici au pus capăt cu înțelepciune ideii unei boli divine, mai întâi recunoscând politicos că zeii cauzează totul și apoi tratând fiecare boală ca o afecțiune nici mai mult, nici mai puțin divină decât oricare alta. Fără a recomanda restricții conceptuale, aș - propune un regim de reținere și grijă atunci când oamenii serioși (fie ei filosofi, oameni de știință sau pur și simplu persoane interesate) iau în considerare amploarea și autoritatea reflecțiilor filosofice și științifice despre *Lebenswelt,* lumea vieții. Se poate și ar trebui să o admirăm *pe Antigona* și să învățăm de la ea, fără a cere apoi includerea ei în codurile juridice pentru soluționarea disputelor în instanțele de judecată . Se poate și ar trebui să admirăm protecția pe care medicii hipocratici s-au străduit să o ofere medicamente împotriva convingerilor religioase ritualice, fără a insista ca rugăciunea să fie interzisă în camerele de urgență. Și se poate și ar trebui să admirăm cercetarea detaliată, repetabilă și riguroasă a comunității neuroștiințifice, fără a dezvolta o atitudine sceptică față de bunul simț și față de recunoașterea altfel insistentă a noastră înșine ca noi înșine .

Indiferent dacă analiza lui Peter Hacker asupra absurdităților conceptuale ale neurocercetătorilor de astăzi este pe deplin corectă sau dacă judecățile ulterioare o

condamnă, nu există nicio îndoială că el demonstrează fidelitate față de misiunea filosofică, demonstrând în același timp stăpânirea resurselor pe care filosofia le-a dezvoltat pentru această misiune. El nu solicită titlul de membru onorific al Societății pentru Neuroștiințe și nici nu intenționează să îmbogățească acea bază de date deja impresionantă pe care se bazează progresul științific. De fapt, descoperirile cu adevărat importante ale neuroștiinței cognitive au fost opera unei mici legiuni de specialiști, în cea mai mare parte necunoscuți cititorilor Times *Literary Supplement / The* New *York Review of Books.* Cei care au plasat știința vederii pe harta științei autentice, de vârf, nu au oferit nicio mare „neurofilosofie". Numele lor nu vor însemna nimic pentru majoritatea cititorilor umilelor mele scrieri: Selig Hecht, M.H. Pirenne, Clarence Graham, H.K. Hartline și George Wald. Pitts și McCulloch au rămas atașați de matematica lor și au propus circuite care, dacă sunt proiectate inteligent, dau rezultate remarcabile. Pavlov a fost productiv când s-a dedicat chimiei digestiei, dar a devenit un fel de scriitor de treabă când a încercat să traducă întreaga psihologie în limbajul „dinamicii creierului". Du Bois a fost mai înțelept când s-a confruntat cu această *întrebare dificilă* și a concluzionat, sunt sigur cu bucurie în voce: *„IGNORABIMUS!"*

Dacă Hacker ar căuta să devină membru al oricărui cerc ortodox în care se cer cotizații sub formă de clișee, ar face-o în *cercul anticartezian.* Prin clișeu, nu înțeleg nimic mai mult decât o expresie sau o maximă banală. „Dumnezeu este bun" este un astfel de clișeu, iar credincioșii vor considera frecvența unei astfel de afirmații ca o măsură a adevărului său. „Resping cartesianismul" este o altă afirmație banală care poate conține, de asemenea, o reflecție asupra unui adevăr mai profund sau mai înalt. Dar cartesianismul înseamnă lucruri diferite pentru diferiți membri ai cercului anticartezian. Într-un mod destul de neglijent, este în general acuzat de o ontologie a două substanțe combinată cu o teorie a „teatrului minții". Apoi, atât dualismul, cât și teatrul interior sunt respinse cu seninătate ca dovezi ale inocenței filosofice.

Presupun că este încă permis să reamintim tuturor că Descartes a fost fondatorul geometriei analitice, un adevărat maestru al științei opticii, care a înțeles pe deplin știința timpului său și care, în scrisorile și corespondența sa, a reunit cele mai fine minți ale unei epoci populate de mari gânditori. Puține dintre criticile

aduse ideilor sale în ultimele două secole nu ar fi fost anticipate de Hobbes, Gassendi și părintele Mersenne, cu care Descartes a avut o dezbatere aprinsă în scris. După cum i-a spus prințesei Elisabeta , a considerat util să adopte în scrierile sale expresii și analogii care erau poate prea filozofice, pentru a evita să fie înțeles greșit. Știind ce știa despre materie, era convins că caracterul esențial al vieții raționale și perceptive nu putea fi derivat din materie în nicio combinație posibilă. Există ceva cu adevărat ridicol într -o ontologie dualistă care contrastează ființele extinse și cele neextinse? Nu cred. De fapt, *nimeni nu crede asta,* pentru că este exclusă de gândire însăși.

Este acesta un argument pentru dualism? În diverse momente, pe o perioadă de poate cincizeci de ani, m-am întrebat câte tipuri diferite de „materie" ar putea constitui totalitatea realității. Am ajuns să înțeleg două tipuri, pe care, în lipsa unui termen mai bun, le voi numi materie fizică și orice stă la baza dimensiunilor morale, estetice, raționale și emoționale ale vieții mele. Oh, haideți să-i spunem „mentală". În realitate, doar o porțiune foarte mică din întregul spectru electromagnetic îmi este vizibilă; doar „materia" care are o lungime de undă între 3.600 și 7.600 angstromi. În cazul în care capacitatea mea de a înțelege totalitatea realității este la fel de limitată ca viziunea mea, cel mai bine este să las fără răspuns întrebările despre câte tipuri distincte de entități alcătuiesc „tot ceea ce există". Nu știu numărul. Dennett nu știe. Hacker nu știe. Dar dacă „fizica este completă"? Oare aceasta rezolvă problema? După cum pare să fi spus un fost președinte american : „Totul depinde de ce înțelegi prin «complet»".

Hacker aparține școlii de filosofie wittgensteiniene, ai cărei principali susținători și critici sunt recrutați dintre cei mai educați filosofi. Întrebarea dacă subiectele filosofice sunt considerate „puzzle-uri" care trebuie puse laolaltă sau „probleme" care trebuie rezolvate este de mare importanță. În ambele privințe, claritatea și coerența exprimării sunt esențiale . Nimeni nu ar putea susține în mod serios că, nefiind wittgensteinian, nu este obligat să analizeze aparatul cultural și lingvistic prin care sunt create și împărtășite conceptele.

Hacker scrie cu precizie, până la punctul de a se apropie de ruda sa etimologică, *prețiozitatea.* El acordă atenție detaliilor. Cititorii ar putea reacționa la o astfel de grijă în același fel în care reacționăm noi la șoferii care nu depășesc

niciodată limita de viteză. Când spune că

> Adevărurile conceptuale definesc spațiul logic în care se situează faptele. Ele determină ce are sens.

S-ar putea interpreta că subestimează faptele sau că acordă filosofiei puterea de a le guverna. El nu face așa ceva. Cosmosul este plin de fapte, a căror imensă pluralitate depășește simțurile și chiar înțelegerea noastră. Din acest foc viu, radiant, extragem câteva fragmente - cele vizibile sau aproape vizibile - și începem să țesem o poveste. În rare ocazii , povestea este atât de sistematică, atât de adaptată la fragmentele de care dispunem, încât alte povești decurg din prima, apoi altele, și în curând ne trezim posedați de puteri practic profetice cu privire la ceea ce va urma. Filosoful este însă cel care trebuie să țină în frâu entuziasmul celor care spun poveștile, deoarece, lăsați de capul lor, ar putea invoca un viitor care nu face decât să valideze confuzia noastră prezentă .

NOTE

INTRODUCERE ÎN *FUNDAMENTELE FILOSOFICE ALE NEUROȘTIINȚEI*

Ceea ce urmează este textul nemodificat al prefeței la *Fundamentele filozofice ale neuroştiinţei* (denumită în continuare PFN), cu excepția reducerii ultimelor două paragrafe și a eliminării referințelor încrucișate, înlocuite cu note acolo unde este necesar.

1. Obiecțiile metodologice la aceste distincții sunt discutate mai jos și mai detaliat în PFN, capitolul 14.
2. Prin urmare, capitolul 1 din PFN începe cu un studiu istoric al primelor progrese în neuroștiințe.
3. Prin urmare, capitolul 2 din PFN este dedicat unei examinări critice a angajamentelor sale conceptuale.
4. Notă de confidențialitate §3.10.
5. Vezi mai jos, extrasul din capitolul 3 al lucrării PFN. Capitolul original este mult mai lung decât extrasul prezentat aici, prin urmare, argumentul este mai detaliat.
6. Reducționismul este examinat în capitolul 13 din PFN.
7. În capitolul 14 din PFN.
8. Vezi Notificarea Prescripției Generale §14.3.
9. Vezi capitolele 1 și 2 din PFN.

10. În PFN §6.31, unde se studiază imageria mentală, și în PFN §8.2, unde se investighează mișcarea voluntară, sunt analizate exemple care demonstrează inutilitatea investigației.

11. Exemple sunt oferite în discuția despre memorie în PFN §§5.21-5.22 și despre emoții și pofte în PFN §7.1.

12. Tratăm în detaliu dificultățile metodologice în capitolul 3, §3 din PFN (în această carte) și în capitolul 14 din PFN.

EXTRAS DIN CAPITOLUL 3 DIN *FUNDAMENTELE FILOZOFICE ALE NEUROȘTIINȚEI*

Aceste pagini reproduc textul nemodificat al paginilor 68-80 din PFN, cu excepția referințelor încrucișate, care, acolo unde a fost necesar, au fost retrogradate în note.

1. F. Crick, *Ipoteza uimitoare* (Londra, Touchstone Books, 1995), pp. 30, 32 și următoarele, 57.

2. G. Edelman, *Aer strălucitor, foc strălucitor* (Londra, Penguin Books, 1994), pp. 109 și urm., 130.

3. C. Blakemore, *Mecanica minții* (Cambridge, Cambridge University - Press, 1977), p. 91.

4. JZ Young, *Programs of the Brain* (Oxford, Oxford University Press, 1978), p. 119 (traducere în spaniolă, *Los programas del cerebro,* Mexic, Fondo de Cultura Econòmica, 1986).

5. A. Damasio, *Eroarea lui Descartes — Emoția, rațiunea și creierul uman* (Londra, Papermac, 1996), p. 173 (Traducere în spaniolă: *El error de Des cartes. La emoción, la razón y el cerebro humano,* Barcelona, Crítica, 1999).

6. B. Libet, „Inițiativa cerebrală inconștientă și rolul voinței conștiente în acțiunea voluntară", *The Behavioral and Brain Sciences* (1985), 8, p. 536.

7. JP Frisby, *Seeing: Illusion, Brain and Mind* (Oxford, Oxford University Press, 1980), pp. 8 și urm. Este curios în acest moment că modul filosofic echivoc

de a vorbi asociat cu tradițiile carteziene și empiriste, vorbind în mod specific despre lumea „externă", a fost transferat de la minte la creier. Era echivoc deoarece pretindea să contrasteze o „lume internă a conștiinței" cu o „lume externă a materiei". Dar acest lucru este confuz. Mintea nu este un fel de loc, iar ceea ce se spune în acest mod de a vorbi că este *în* minte nu este prin urmare localizat spațial. Prin urmare, nici lumea (care nu este „simplă materie", ci și ființe vii) nu este *spațial* „în afara" minții. Contrastul dintre ceea ce este în creier și ceea ce este în afara lui este, desigur, perfect literal și incontestabil. Ceea ce este discutabil este afirmația că există „descrieri simbolice" în creier.

8. R. L. Gregory, „Ochiul încurcat", în RL Gregory și EH Gombric (comp.), *Iluzie în natură și artă* (Londra, Duckworth, 1973), p. 50.

9. D. Marr, *Viziunea, o investigație computațională asupra reprezentării și procesării informațiilor vizuale la om* (San Francisco, Freeman, 1980), pag. 3.

10. P.N. Johnson-Laird, „Cum ar putea conștiința să apară din calculele creierului?", în C. Blakemore și S. Greenfield (comp.) *Mindwaves* (Oxford, Blackwell, 1987), p. 257.

11. Susan Greenfield, explicând publicului său realizările tomografiei cu emisie de pozitroni, anunță cu admirație că pentru prima dată este posibil să vedem *gândurile.* Semir Zeki îi informează pe membrii Societății Regale că noul mileniu aparține neurobiologiei, care, printre altele, va rezolva problemele străvechi ale filosofiei (vezi S. Zeki, „Splendori și mizerii ale creierului", *Phil. Trans. R Soc. Land.* B (1999), 354, 2054. Vezi PFN §14.42).

12. L. Wittgenstein, *Investigații filosofice* (Oxford, Blackwell, 1953), §281 (traducere în spaniolă: *Investigaciones filosóficas,* Barcelona, Crítica, 1988) (vezi și §§282-284, 357-361). Reflecția fundamentală asupra acestei observații a fost făcută de A.J.P. Kenny, „The Homunculus Fallacy" (1971), retipărită în lucrarea sa *The Legacy of Wittgenstein* (Oxford, Blackwell, 1984), pp. 126-136. Pentru o interpretare detaliată a observației lui Wittgenstein, vezi P.M.S. Hacker, *Wittgenstein: Meaning and Mind, volumul 3 al unui comentariu analitic asupra investigațiilor filosofice* (Oxford, Blackwell, 1990), Exegesis §§ 281-184, 357-361 și eseul intitulat „Oameni, minți și mașini", care examinează unele dintre ramificațiile ideii lui Wittgenstein. După cum se demonstrează clar în capitolul 1

al PFN, Aristotel l-a anticipat în această idee (DA 408 [b] 2-15).

13. Kenny *(ibid.,* (p. 125) folosește expresia „eroarea homunculului ” pentru a se referi la eroarea conceptuală în cauză. Deși este o expresie foarte grafică, autorul admite că poate fi confuză, deoarece eroarea *nu este* pur și simplu aceea de a atribui predicate psihologice unui homuncul imaginar care se află în cap. În opinia noastră, expresia „eroare mereologică” este mai potrivită. Trebuie menționat, totuși, că eroarea în cauză nu este doar eroarea de a atribui unei părți predicate care se aplică doar unui întreg, ci mai degrabă un caz special al acestei confuzii mai generale. După cum subliniază Kenny, aplicarea eronată a unui predicat nu este, strict vorbind, o eroare, deoarece nu este o formă de raționament invalid, dar duce la erori *(ibid.,* pp. 125 și următoarele). Nu există nicio îndoială că această confuzie mereologică este comună atât în rândul psihologilor, cât și al neurocercetătorilor.

14. Principii mereologice comparabile se aplică obiectelor neînsuflețite și unora dintre proprietățile acestora. Faptul că o mașină este rapidă nu implică faptul că carburatorul său este rapid, iar faptul că un ceas indică ora cu precizie nu implică faptul că balansierul său indică ora cu exactitate.

15. Să subliniem, totuși, că atunci când mă doare mâna, eu sunt cel care suferă durerea, nu mâna mea. Și când cineva îmi doare mâna, mă doare pe mine. Verbele de senzație (spre deosebire de cele de percepție) se aplică părților corpului; adică, corpul nostru este sensibil și părțile sale pot suferi, pot mânca etc. Însă expresiile verbale corespunzătoare care încorporează sintagme nominale, de exemplu, „a avea durere (mâncărime etc.)”, sunt predicabile doar persoanei respective, nu părților corpului acesteia (unde se află senzația).

16. Vezi Simon Ullman, „Tacit Assumptions in the Computational Study of Vision”, în A. Gotea (ed.), *Representations of Vision, Trends and Tacit Assumptions in Vision Research* (Cambridge, Cambridge University Press, 1991). Autorul își limitează discuția la utilizarea (sau, în opinia noastră, la utilizarea greșită) a unor termeni precum „reprezentare” și „reprezentare simbolică”.

17. Expresia este a lui Richard Gregory; vezi „The Confounded Eye” în RL Gregory și EH Gombrich (eds.), *Illusion in Nature and Art* (Londra, Duckworth, 1973), p. 51.

18. Vezi C. Blakemore, «Înțelegerea imaginilor în creier», în H. Barlow, C. Blakemore și M. Wetson-Smith (ed.), *Images and Understanding* (Cambridge, Cambridge University Press, 1990), pp. 257-283 (traducere în spaniolă: *Imagen y conocimiento,* Madrid, Crítica, 1994).

19. S. Zeki, „Abstracție și idealism", *Nature,* 404 (aprilie 2000), p. 547.

20. JZ Young, *Programs of the Brain (* Oxford, Oxford University Press, 1978), p. 192 (traducere în spaniolă, Los programas del cerebro, Mexic, Fondo de Cultura *Económica* , 1986).

21. Brenda Milner, Larry Squire și Eric Kandel, „Neuroștiința cognitivă și studiul memoriei", *Neuron,* 20 (1998), p. 450.

22. Pentru o expunere detaliată a acestei afirmații discutabile, consultați PFN §5.22.

23. Ullmann, *ibid.,* p. 314 și urm.

24. Marr, *ibid.,* p. 20.

25. Marr, *ibid.,* p. 21.

26. Marr, *ibid.*

T I. Pentru critici suplimentare la adresa expunerii computaționale a vederii de către Mart, vezi PFN §4.24.

28. Frisby, *ibid.,* p. 8.

29. Roger Sperry, « Specializare laterală în emisferele separate chirurgical », în FO Schmitt și FG Worden (eds.), *The Neurosciences Third Study Programme* (MIT Press, Cambridge, Mass., 1974), p. II. Pentru o analiză detaliată a acestor forme de descriere, vezi PFN §14.3.

30. Blakemore, „Înțelegerea imaginilor din creier", p. 265. Trebuie menționat că ceea ce este necesar pentru a recunoaște ordinea creierului *nu este* un set de *reguli,* ci pur și simplu un set de corelații regulate. O regulă, spre deosebire de simpla regularitate, este un criteriu de conduită, un standard de corectitudine după care se poate judeca dacă o conduită este corectă sau greșită, adecvată sau nepotrivită.

31. JZ Young, *Programele creierului* (Oxford University Press, Oxford , 1978), p. 52. (trad., distribuție, *Programele creierului,* Mexic, Fon do de Cultura Economică, 1986).

32. Blakemore, *ibid.,* pp. 265-2

33. JZ Young, *Programele creierului,* p. II.

34. Confuzia generată de incapacitatea de a distinge o regulă de o regularitate și normativul de cauzal este evidentă în comentariile lui Blakemore la diagrama motorică a „homunculului" lui Penfield și Rasmussen. Blakemore insistă că „maxilarele și mâinile sunt atât de suprareprezentate" („Înțelegerea imaginilor din creier", p. 266, în nota explicativă extinsă de la Fig. 17.6); dar acest lucru ar avea sens doar dacă am vorbi despre o hartă cu o metodă de proiecție înșelătoare - în acest sens, vorbim despre distorsiunile relative ale proiecției Mercator (cilindrice). Dar, din moment ce tot ce reprezintă desenele este numărul relativ de celule responsabile cauzal pentru anumite funcții, *nimic nu* este sau nu ar putea fi suprareprezentat. Cu siguranță, Blakemore nu înseamnă că există mai multe celule în creier corelate cauzal cu maxilarele și mâinile decât ar trebui să existe.

EXTRAS DIN CAPITOLUL 10 DIN *FUNDAMENTELE FILOZOFICE ALE NEUROȘTIINȚEI*

1. Ned Block, «Qualia», în S. Guttenplan (comp.), *Blackwell Companion to the Philosophy of Mind* (Blackwell, Oxford, 1994), p. 514.

2. R. Searle, „Conștiință", *Revista Anuală,* p. 560.

3. Searle, *ibid.,* p. 561.

4. Chalmers, *Mintea conștientă* (Oxford University Press, Oxford, 1996), p. 4.

5. DJ Chalmers, *Mintea conștientă,* p. 10.

6.1 . Glynn, *Anatomia gândirii,* p. 392.

7. A. Damasio, *Sentimentul a ceea ce se întâmplă,* p. 9. Rețineți că în acest moment se presupune, fără nicio demonstrație, că culoarea și sunetul nu sunt proprietăți ale obiectelor, ci ale impresiilor senzoriale.

8. G. Edelman şi G. Tononi, *Consciousness — How Matter Becomes - Imagination,* p. 157.

9. E. Lomand, „Conștiință”, în *Enciclopedia de Filosofie Routledge (* Routledge, Londra, 1998), vol. 2, p. 581.

10. Searle, *Misterul conștiinței,* p. XIV.

11. T. Nagel, „Cum e să fii liliac?”, retipărire în *Întrebări mortale* (Cambridge University Press, Cambridge, 1979), p. 166.

12. Nagel, *ibid.,* p. 170n.

13. Nagel, *ibid.,* p. 170.

14. M. Davies și GW Humphreys (ed.), *Conștiință* (Blackwell, Oxford, 1993), p. 9.

15. Edelman și Tononi, *Conștiința - Cum materia devine imaginație ,* p. 11.

16. Chalmers, *Mintea conștientă,* p. 4.

17. Vezi Searle, *Misterele conștiinței,* p. 201.

FRAGMENT DIN CAPITOLUL 14 DIN *FUNDAMENTELE FILOZOFICE ALE NEUROȘTIINȚEI*

1. Vezi, de exemplu, discuția despre mișcările voluntare din PFN §8.2.
2. Vezi, de exemplu, discuția despre imageria mentală din PFN §14.3.

3. Vezi, de exemplu, PFN, §14.3.
4. Așa cum se menționează în PFN §4.23.
5. Așa cum se menționează în PFN §§6.3-6.31.
6. Vezi PFN §2.3.

NEUROȘTIINȚĂ ȘI FILOSOFIE

1. Această modalitate de utilizare a termenilor precum „a stoca” și „memorie” a fost ținta unor critici semnificative: vezi PFN, pp. 158-171.

2. Profesorul Dennett notează în nota de subsol 15 că, în timpul reuniunii anuale a APA, Bennett și-a exprimat „profunda consternare față de modelele și

ipotezele senzaționaliste ale neurocercetătorilor cognitivi actuali și a precizat că, în opinia sa, toate erau de neînțeles. Cu un informator precum Bennett, nu este de mirare că Hacker nu a reușit să găsească nimic valoros în neuroștiința cognitivă". De asemenea, el sugerează că sunt prins iremediabil în acea „animositate reciprocă" care predomină între neurocercetătorii sinaptici și cei cognitivi. Nu este cazul. În primul rând, David Marr este considerat un neurocercetător cognitiv talentat în manualele pe această temă (vezi Gazzaniga, Ivry și Mangun, 2002, p. 597); am publicat articole despre teoria rețelelor sinaptice în spiritul operei lui Marr și nu consider acest lucru în niciun fel o demonstrație de ostilitate nechibzuită față de neuroștiința cognitivă (vezi, de exemplu, Bennett, Gibson și Robinson, 1994). O carte viitoare scrisă de Hacker și de mine, * *Istoria neuroștiinței cognitive**, nu ar fi fost scrisă dacă am fi fost prinși fără speranță într-o ostilitate irațională față de neuroștiința cognitivă. În al doilea rând, la întâlnirea APA, nu am spus că „modelele neurocercetătoarelor cognitive actuale " erau „toate de neînțeles". Ceea ce am făcut a fost să subliniez complexitatea extremă a biologiei pe care o modelează și insuficiența consecventă a cunoștințelor noastre biologice. Acest lucru face foarte dificilă construirea de modele care să clarifice funcțiile rețelelor sinaptice . Exemple invocate în susținerea acestui punct sunt furnizate în a doua secțiune a acestui capitol . Cu toate acestea, la întâlnirea APA am adăugat că este ciudat să spunem despre astfel de rețele și grupuri de rețele că „văd", „își amintesc" etc., adică că posedă atributele psihologice ale ființelor umane (vezi secțiunea 3).

FILOSOFIA CA ANTROPOLOGIE NAIVĂ

1. Scopul meu în lucrarea *„Conținut și conștiință"* (traducere în spaniolă: *Contenido y conciencia,* Barcelona, Gedisa, 1996) din 1969 a fost „să stabilesc cadrul conceptual din care trebuie spusă întreaga poveste, să determin restricțiile care sunt impuse oricărei teorii satisfăcătoare (p. ix) [...] [să dezvolt] ideea unui mod distinct de discurs, limbajul minții , pe care îl folosim în mod normal pentru a descrie și explica experiențele noastre mentale și care poate fi legat doar indirect

de modul de discurs în care este formulată știința" (p. x).

2. Deși teoreticienii anteriori - Freud, de exemplu - au vorbit despre *psihologia populară într* -un sens oarecum diferit, cred că am fost primul, în „Trei tipuri de psihologie intențională" (1978), care a propus utilizarea ei pentru a descrie ceea ce Hacker și Bennett numesc „descriere psihologică actuală ". Ei, la fel ca mine, insistă că nu este o *teorie.*

3. Vezi discuția mea despre acest punct în „A Cure for the Common Code", în *Brainstorms* (1978) și, mai recent, în „Intentional Laws and Computational Psychology" (secțiunea 5 din „Back from the Drawing Board") în Dahlbom (ed.), *Dennett and His Critics,* 1993.

4. Lista este lungă. Vedeți criticile mele, pe lângă lucrările citate în notele precedente, la adresa lucrărilor despre imagistică, *qualia,* introspecție și durere din *Brainstorms.* Nu sunt singurul teoretician a cărui lucrare nu o citează, în ciuda faptului că aceasta anticipează propria lor lucrare. De exemplu, în explicația lor despre imagistica mentală, ei reinventează în mod accidental mai multe dintre ideile lui Zenon Pylyshyn. Bennett și Hacker nu sunt primii analiști conceptuali care frecventează aceste ape, iar majoritatea, dacă nu toate, punctele lor de vedere au fost exprimate anterior și examinate în mod corespunzător în lucrări pe care nu le citează. Nu am găsit nimic nou în cartea lor.

5. Anexa pe care o dedică atacării ideilor mele este un comentariu lung și disprețuitor, o colecție de interpretări stupid de eronate care se încheie astfel: „Dacă argumentele noastre sunt valide, atunci teoriile lui Dennett despre intenționalitate și conștiință nu contribuie cu nimic la clarificarea filosofică a intenționalității sau conștiinței. Nici nu oferă vreo îndrumare pentru cercetarea în neuroștiințe sau înțelegerea neuroștiinței" (p. 435). Dar nu există argumente, ci doar afirmații de „incoerență". La reuniunea anuală APA, unde a fost prezentat acest eseu, Hacker a răspuns cu mai mult din același lucru. La Oxford, în anii 1960, o ridicare delicată din umeri era adesea confundată cu un argument. Acele zile au apus. Sfatul meu pentru Hacker: dacă găsești aceste subiecte de neînțeles, încearcă să le studiezi mai sârguincios. Abia ai început pregătirea ta în științe cognitive.

6. Ignorarea de către Hacker a distincției mele nu poate fi atribuită miopiei;

pe lângă Hornsby, aceasta a fost discutată în profunzime de alți filozofi de la Oxford, de exemplu, Davies, 2000; Hurley, *Synthèse,* 2001; și Bermudez, «Nonconceptual Content: From Perceptual Experience to Subpersonal Computational States», *Mind and Language,* 1995.

7. Vezi și „Condițiile personalității" în *Brainstorming.*

8. Vezi și expunerea nivelurilor de explicație în *Science Explained* (1991).

9. Searle, care a prezentat o serie de obiecții la adresa lui Bennett și Hacker, abia a menționat această problemă la întâlnirea APA unde a fost prezentat acest eseu .

10. Pentru un filosof care nu consideră adevărul și falsitatea drept pietre de temelie ale propozițiilor filosofice, Hacker este extrem de liberal în afirmațiile sale grosolane, fără a argumenta că un lucru este interpretat greșit , altul este o greșeală și așa mai departe. Aceste remarci trecătoare sunt dificil de interpretat fără a presupune că sunt menite să fie adevărate (spre deosebire de false). Poate că ar trebui să înțelegem că doar o mică parte din propozițiile lui Hacker, cele specifice filosofice , „preced" adevărul și falsitatea, în timp ce marea majoritate a propozițiilor sale sunt ceea ce par a fi: afirmații care indică adevărul. Și, ca atare, probabil, ele sunt supuse confirmării și infirmării empirice .

11. În *lucrarea „Vise dulci: Obstacole filozofice în calea unei științe a conștiinței" (* 2005) (trad. Katz Barpal, 2006), descriu unele tensiuni din filosofia contemporană a minții ca *antropologie naivă a priori a sinelui (pp. 31-35). Opera lui Hacker mi se pare un* caz paradigmatic .

12. Rețineți că nu spun că autoantropologia este întotdeauna o întreprindere prostească sau inutilă. Ceea ce spun este că este o cercetare empirică ce produce rezultate - atunci când este condusă corect - cu privire la intuițiile pe care cercetătorii le descoperă în ei înșiși și la implicațiile acestor intuiții. Aceste rezultate pot fi rodnice pentru cercetare, dar rămâne de văzut în ce condiții aceste implicații pot fi considerate în mod serios ca ghiduri către adevăr pe orice subiect. Pentru mai multe detalii despre acest subiect, consultați *„Vise dulci".*

13. *Revendicarea rațiunii: Wittgenstein, scepticism, moralitate și tragedie* (1979, ediția a 2-a , 1999) (trad., distribuție: *Revendicări ale rațiunii,* Madrid, Sínthesis , 2003).

14. Este posibil ca un filosof precum Hacker *să aibă dreptate* chiar și atunci când scopul său nu este adevărul?

15. Este rezonabil să presupunem că Bennett, distinsul neurocercetător, a jucat rolul de informator pentru antropologul Hacker, dar cum să explicăm atunci insensibilitatea aproape completă a lui Hacker față de subtilitățile *jargonului* științei cognitive (și modelele și descoperirile) ? A ales Hacker informatorul greșit? Poate că cercetările neuroștiințifice ale lui Bennett au fost efectuate la nivelul sinapsei, iar cei care lucrează la acest nivel subneural sunt la fel de departe de disciplinele științei cognitive precum biologii moleculari sunt de etologii de teren. Nu există prea multă comunicare între astfel de întreprinderi îndepărtate și, chiar și în cele mai bune circumstanțe, există multă neînțelegere - și, din păcate, o mare animozitate reciprocă. Îmi amintesc de un distins director de laborator care a început un atelier cu următoarea observație: „În laboratorul nostru, avem o vorbă: dacă lucrezi la un neuron, este neuroștiință; dacă lucrezi la doi neuroni, este psihologie." Nu a spus-o ca pe un compliment . Alegerea unui informator ostil este, desigur, o rețetă pentru dezastrul antropologic. (Adăugat după întâlnirea APA.) Bennett a confirmat această conjectură în discursul său de deschidere; după ce și-a analizat cariera de cercetător al sinapselor, și-a exprimat profunda consternare față de modelele și ipotezele senzaționaliste ale neurocercetărilor cognitive actuale și a precizat că, în opinia sa, toate erau de neînțeles. Cu un informator precum Bennett, nu e de mirare că Hacker nu a putut găsi nimic valoros în neuroștiința cognitivă .

16. Vezi lucrarea mea **Content and Consciousness**,* p. 183 (traducere în spaniolă: **Content and Consciousness**,* Barcelona, Gedisa, 1996).

17. Ca să dau doar un exemplu, atunci când Hacker deplânge „ locuțiunea mea verbală barbară «a fi o chestiune de»" (p. 422) și insistă că „receptorul opioid *este la fel de mult o chestiune* de opioide precum pisicile sunt de câini sau vulpile de vulpi" (p. 423), are, desigur, perfectă dreptate: relația elegantă dintre opioide și receptorul opioid nu este o „chestiune de" propriu-zisă (îmi cer scuze pentru termenii întortocheați), ci o simplă „proto-chestiune de" (iată-ne din nou!), dar acesta este *exact* genul de proprietate pe care am putea-o aprecia într-o simplă parte a unei anumite sume logice care (organizată corespunzător) ar putea

demonstra o intenționalitate bona fide, autentică, filosofic solidă, paradigmatică...

18. În sensul restrâns de Hacker.

19. Aceasta a fost o temă recurentă în lucrările critice despre științele cognitive . Articolele clasice datează de la „Ce se află într-o legătură" de William Wood (în Bobrow și Collins, *Representation and Understanding,* 1975) și „Inteligența artificială întâlnește prostia naturală" de Drew McDermott, în Haugeland, *Mind Design (1981), și includ Cognition and Reality* (1975) de Neisser (*traducere în spaniolă: Procesos conozco y realidad,* Madrid, Marova, 1981) și „Inteligență fără reprezentare" de Rodney Brook, *Inteligență artificială* (1991). Lista continuă și în ziua de azi, incluzând contribuții ale unor filosofi care și-au făcut temele și sunt familiarizați îndeaproape cu aceste subiecte.

20. Hacker și Bennett spun: „Ar fi o greșeală, deși inofensivă, să vorbim despre hărți cerebrale când ceea ce vrem să spunem este că anumite caracteristici ale câmpului vizual pot fi mapate pe descărcările grupurilor de celule din cortexul vizual primar sau striat. Dar de aici nu se poate continua spunând, așa cum face Young, că creierul își folosește hărțile atunci când formulează ipoteze despre ceea ce este vizibil" (p. 77). Dar acesta este motivul pentru care discuțiile despre hărți sunt perspicace: faptul că creierul le folosește *ca hărți* . Altfel, este clar că acest mod de a vorbi ar fi lipsit de sens. Și acesta este motivul pentru care accentul pus de Kosslyn pe modelele vizibile de excitație din cortex în timpul imagisticii este complet irelevant în ceea ce privește natura proceselor care stau la baza a ceea ce numim, la nivel personal, imagistică vizuală. Vezi articolul principal al lui Pylyshyn din numărul din aprilie 2002 al *BBS* și comentariul meu, *„Utilizează creierul tău imaginile de pe el și, dacă da, cum?"*.

21. „Filosofii nu ar trebui să se afle în situația de a abandona teoriile lor preferate despre natura conștiinței în fața dovezilor provenite din rezultatele științifice. Nu ar trebui să aibă teorii preferate , pentru că, în primul rând, nu ar trebui să postuleze teorii empirice supuse confirmării și infirmării empirice. Treaba lor sunt conceptele, nu judecățile empirice; formele gândirii, nu conținutul acesteia; ceea ce este logic posibil, nu ceea ce este empiric real; ceea ce are sens și ceea ce nu, nu ceea ce este și nu este adevărat" (p. 404). Această îngustă vedere în ceea ce privește propria treabă a filosofului este ceea ce îl face pe Hacker să

greșească atât de spectaculos atunci când încearcă să critice oamenii de știință.

22. Pentru un exemplu de acest *tip* de explicație, consultați relatarea mea simplificată despre cum robotul Shakey distinge cutiile de piramide (un talent „la nivel personal" al unui robot) prin trasarea (la nivel subpersonal) a unor desene liniare ale imaginilor sale retiniene și apoi utilizarea programului său de semantică liniară pentru a identifica trăsăturile distinctive ale cutiilor, în *Consciousness Explained.*

23. „Anexa I: Daniel Dennett" a lui Bennett și Hacker nu merită o respingere detaliată, având în vedere interpretările greșite frecvente ale pasajelor citate în afara contextului și omiterea aparent deliberată a oricărei discuții despre pasajele în care mă apăr în mod specific împotriva interpretărilor greșite pe care le fac în mod repetat, așa cum am subliniat deja. Cu toate acestea, nu pot rezista tentației de a observa că sunt induși în eroare de eroarea creaționistă de a anticipa orice explicație a caracteristicilor biologice în termenii a ceea ce eu numesc atitudinea designului : „Evoluția nu a *proiectat* nimic - realizarea lui Darwin a fost să înlocuiască explicația în termeni de design cu explicații evolutive" (p. 425). Se pare că ei nu înțeleg cum funcționează explicația evolutivă.

PENTRU A READUCE CONȘTIINȚA ÎN CREIER

Le sunt îndatorată Romeliei Drager, Jenniferi Hudin și Dagmar Searle pentru comentariile lor la primele versiuni ale acestui articol.

1. De exemplu, John R. Searle, *Redescoperirea minții,* Cambridge , MIT Press, 1992 (traducere în spaniolă: *Redescoperirea minții,* Barcelona , Critica, 1996).

2. Pentru posibile contraexemple la această afirmație, vezi, de exemplu, explicația lui Christof Koch despre „neuronul lui Halle Berry", *New York Times,* 5 iulie 2005.

3. John R. Searle, *Raționalitatea în acțiune* (Cambridge, MIT Press, 2001).

Ipotezele conceptuale ale neuroștiinței cognitive

1. MR Bennett și PMS Hacker, *Fundamentele filozofice ale neuroștiinței (* Oxford, Blackwell, 2003); referințele la această carte se vor face folosind inițialele PFN.

2. Profesorul Searle afirmă că un rezultat conceptual este important doar ca parte a unei teorii generale (p. 122). Dacă prin „o teorie generală ” se referă la o explicație generală a unei rețele conceptuale și nu doar la rezultate izolate, suntem de acord cu el. Când negăm că explicațiile noastre generale sunt teoretice, vrem să spunem că nu se află la același nivel logic ca teoriile științifice. Sunt descrieri , nu ipoteze; nu pot fi confirmate sau infirmate prin experiment ; nu sunt ipotetico-deductive și scopul lor nu este nici predicția , nici furnizarea de explicații cauzale; nu implică idealizări în sensul în care o face știința (de exemplu, ideea de masă a unui punct în mecanica newtoniană) și nu aproximează fapte empirice în limitele unor marje de eroare convenite; nu caută să descopere entități noi și nici nu introduc entități ipotetice în scopuri explicative.

3. Se pare că profesorul Dennett a avut probleme cu această idee. În criticile sale (p. 79), a citat selectiv din cartea noastră: „ Întrebările conceptuale preced întrebările despre adevăr și falsitate…” (PFN, p. 2, vezi p. 4 din această carte), „Adevărul și falsitatea sunt pentru știință ceea ce rațiunea și nonsensul sunt pentru filosofie” (PFN, p. 6, vezi p. 12 din această carte). Din aceasta, el a concluzionat că, din perspectiva noastră, filosofia nu este deloc interesată de adevăr. Cu toate acestea, a omis continuarea primei propoziții:

> Acestea sunt întrebări care privesc formele noastre de reprezentare, nu adevărul sau falsitatea *afirmațiilor empirice.* Aceste forme sunt presupuse în afirmațiile științifice adevărate (și false) și în teoriile științifice corecte (și incorecte). Ele nu determină ce este empiric adevărat sau fals, ci mai degrabă ce are sens sau nu.
>
> (PFN, p. 2, vezi p. 4 din această carte; italicele sunt ale noastre)

Profesorul Dennett a omis și observația de pe pagina alăturată , conform

căreia neuroștiința descoperă multe lucruri despre baza neuronală a capacităților umane, „dar descoperirile sale nu afectează în niciun fel *adevărul conceptual* conform căruia aceste capacități și exercitarea lor [...] sunt atribute ale ființelor umane, nu ale părților lor" (PFN, p. 3, vezi p. 6 din această carte; sublinierea este a noastră). Evident, concepția noastră este că filosofia se preocupă de adevărurile conceptuale și că acestea determină ce are sens și ce nu.

4. Profesorul Paul Churchland îndreaptă următoarea observație împotriva concepției noastre: „Încă de la Quine, cea mai mare parte a profesiei filosofice a fost înclinată să spună «nu»" sugestiei că există „adevăruri necesare , constitutive ale semnificațiilor, situate pentru totdeauna dincolo de orice infirmare empirică sau factuală". „Curățarea științei", *Inquiry,* 48 (2005), p. 474. Este îndoielnic că a realizat o anchetă sociologică (oare majoritatea filosofilor cred cu adevărat că adevărurile aritmeticii sunt supuse respingerii empirice, împreună cu orice teorie empirică în care acestea sunt integrate?), și suntem surprinși că un filosof ar putea crede că numărul de locuitori este un criteriu al adevărului.

5. Pentru o critică canonică a afirmațiilor lui Quine despre analiticitate, vezi P. E. Strawson și H. P. Grice, „In Defense of a Dogma", *Philosophical Review,* 1956. Pentru o critică mai recentă și mai meticuloasă a poziției generale a lui Quine, vezi H.-J. Glock, *Quine and Davidson on Language, Thought, and Reality* (Cambridge, Cambridge University Press, 2003). Despre contrastele dintre Quine și Wittgenstein, vezi P. M. S. Hacker, *Wittgenstein's Place in Twentieth-Century Analytic Philosophy* (Oxford, Blackwell, 1996), cap. 7.

6. S-ar putea crede (așa cum a sugerat profesorul Churchland) că ideea lui Descartes conform căreia mintea poate afecta cauzal mișcarea corpului (înțeleasă, conform profesorului Churchland, ca o afirmație conceptuală) este infirmată de legea conservării impulsului . Aceasta ar fi o greșeală. Afirmația carteziană (fie ea conceptuală, fie empirică) ar putea fi infirmată doar dacă ar avea sens ; dar, în absența unui criteriu de identitate pentru substanțele imateriale , nu are. Însăși ideea că mintea este o *substanță* de un anumit fel este incoerentă. Prin urmare, afirmația că mintea, astfel înțeleasă, posedă puteri cauzale nu este nici inteligibilă, nici , *a fortiori* , confirmabilă și nici infirmabilă prin observație și testare experimentală. (Luați în considerare ce rezultat experimental ar putea *fi*

considerat dovadă că este adevărat.)

7. concepție *epistemică* stă la baza atacului amplu al profesorului Timothy Williamson asupra ideii însăși de adevăr conceptual, „Adevărul conceptual", *Proceedings of the Aristotelian Society,* supliment vol. 80 (2006). Concepția pe care o schițează nu este ceea ce mulți mari gânditori, de la Kant până în prezent, înțeleg printr-un „adevăr conceptual". După ce a criticat, spre propria sa satisfacție, concepția epistemică pe care el însuși tocmai a schițat-o, profesorul Williamson concluzionează că *nu există adevăruri conceptuale de niciun fel.* Dar această concluzie nu rezultă deloc. Căci tot ce a demonstrat (în cel mai bun caz) este că nu există adevăruri conceptuale care să se încadreze în patul epistemic procrustian pe care l-a conceput.

8. Punctele aristotelice și anticarteziene pe care le evidențiem sunt: 1) principiul lui Aristotel, pe care îl vom discuta mai jos; 2) identificarea de către Aristotel a *psihicului* cu o serie de capacități; 3) faptul că aceste capacități sunt identificate prin ceea ce permit cuiva să facă; 4) faptul că, pentru a determina dacă o creatură posedă o capacitate, trebuie să-i observăm activitățile ; 5) observația lui Aristotel conform căreia a ne întreba dacă *psihicul* și corpul sunt unul sau două lucruri este o întrebare incoerentă.

9. Nu este, desigur, o eroare de rațiune în sens strict, dar duce la erori de rațiune: inferențe invalide și argumente false.

10. A.J.P. Kenny, „Eroarea homunculului", în M. Grene (ed.), *Interpretations of Life and Mind* (Londra, Routledge, 1971). Preferăm denumirea mai puțin grafică, dar mai precisă, de „eroare mereologică " (și, corelativ, „principiul mereologic"). Am constatat că neurocercetătorii erau înclinați să respingă ca fiind puerilă eroarea de a presupune că un homuncul se află în creier și apoi de a atribui direct atribute psihologice creierului.

11. Evident, nu *cu* creierul său, așa cum face cineva lucruri *cu* mâinile sau ochii, nici cu abilitățile sale. De fapt, nu ar putea face niciunul dintre aceste lucruri dacă creierul său nu ar funcționa normal.

12. D. Dennett, *Conținut și conștiință* (Londra, Routledge și Kegan Paul, 1969) (traducere în spaniolă: *Conținut și conștiință,* Barcelona, Gedisa, 1996).

13. Am fost destul de surprinși să vedem că profesorul Dennett a afirmat că

„principalele sale puncte de dezacord" sunt că nu crede că „nivelul personal de explicație este *singurul* nivel de explicație atunci când vine vorba de mintea și acțiunea umană" și că sarcina de a corela aceste două niveluri de explicație „nu este în afara domeniului filosofului" (p. 79). Nu există niciun dezacord în *această* privință. Oricine a luat vreodată o aspirină pentru a-și ameliora durerea de cap sau care a băut suficient alcool pentru a se simți vesel, certăreț sau taciturn și dorește o explicație a succesiunii evenimentelor , trebuie, fără îndoială, să împărtășească primul punct al lui Dennett. Oricine , la fel ca noi în cele 452 de pagini ale cărții *„Fundamentele filosofice ale neuroștiinței",* s-a dedicat clarificării relațiilor logice dintre conceptele psihologice și neurologice și dintre fenomenele pe care le exprimă, va împărtăși și al doilea punct.

14. L. Wittgenstein, *Investigații filosofice* (Oxford, Blackwell, 1953), §281 (traducere în spaniolă: *Investigații filosofice,* Barcelona, Crítica, 1988).

15. Concepția cartesiană despre corpul uman este greșită în multe privințe. Descartes și-a conceput corpul ca pe o mașină insensibilă - o substanță materială fără senzație. Însă concepția noastră autentică despre corp vorbește despre verbe ale senzației despre corpul pe care îl avem - corpul nostru este cel care ne doare peste tot sau ne mănâncă atât de tare încât nu-l putem suporta.

16. Creierul uman face parte dintr-o ființă umană. Se poate spune, de asemenea, că face parte din corpul despre care se spune că o ființă umană îl are. Cu toate acestea, este semnificativ faptul că ezităm să spunem despre o persoană vie, spre deosebire de un cadavru, că corpul său *are* două picioare sau despre cineva căruia i s-a amputat un membru, că corpul său *are* un singur picior . Posesivul înșelător se aplică ființei umane și unui cadavru uman, dar nu, sau doar ezitant, corpului despre care se spune că o ființă umană vie îl are. Deși creierul face parte din corpul uman, nu există nicio îndoială că nu am spune „corpul meu *are* un creier" sau „creierul corpului meu are meningită". Și acest lucru nu este întâmplător.

17. Suntem de acord cu profesorul Searle că întrebarea despre care animale inferioare sunt conștiente nu poate fi rezolvată prin „ analiză lingvistică" (p. 104). Dar el presupune că ar putea fi rezolvată prin investigarea sistemelor lor nervoase, în timp ce noi subliniem că ar putea fi rezolvată prin studierea comportamentului

animalului în circumstanțele vieții sale. Așa cum stabilim dacă un animal vede referindu-ne la reacțiile sale la ceea ce este vizibil, verificăm și dacă un animal este capabil de conștiință studiind repertoriul său comportamental și reacțiile sale la mediul său. (Acest lucru nu implică faptul că a fi conștient înseamnă a te comporta într-un anumit fel , ci doar că criteriile pentru a fi conștient sunt comportamentale.)

18. Justificarea aplicării predicatelor psihologice altora se bazează pe dovezile disponibile. Aceste dovezi pot fi inductive sau constitutive (criterii). Dovezile inductive, în aceste cazuri, presupun motive sau criterii non-inductive. Criteriile pentru aplicarea unui predicat psihologic rezidă în comportament (nu doar în mișcările corpului) în circumstanțe adecvate. Aceste criterii sunt corectabile. Faptul că anumite motive justifică atribuirea unui predicat psihologic cuiva este parțial constitutiv pentru semnificația predicatului, dar nu-i epuizează semnificația. Criteriile pentru aplicarea unui astfel de predicat sunt distincte de condițiile sale de adevăr - un animal poate experimenta durere și să nu o demonstreze sau să prezinte un comportament dureros fără să o sufere efectiv - (nu suntem behavioristi). Condițiile de adevăr ale unei propoziții care atribuie un predicat psihologic unei ființe sunt distincte de adevărul propoziției în sine. Atât criteriile, cât și condițiile de adevăr sunt distincte de condițiile generale în care activitățile de aplicare sau negare a predicatului creaturilor pot fi desfășurate în mod semnificativ. Însă este o greșeală să presupunem că o condiție a „jocului lingvistic desfășurat" (în cuvintele profesorului Searle) este *apariția* unui comportament observabil public. Căci jocul lingvistic cu predicat psihologic se desfășoară atât cu negația sa, cât și cu afirmarea sa. De asemenea, ar fi o greșeală să confundăm condițiile pentru învățarea unui joc lingvistic cu cele necesare - pentru a-l juca.

19. J.Z. Young, * *Programs of the Brain** (Oxford University Press, 1978), p. 192 (traducere în spaniolă: **Los programas del cerebro humano*,* Mexic, Fondo de Cultura Económica, 1986). Profesorul Dennett sugerează, de asemenea (p. 90) că îl denaturăm pe Crick atunci când susținem că, deoarece spune că creierul nostru creează lucruri și face reprezentări bazate pe experiența sau informațiile sale anterioare (E. Crick, * *The Astonishing Hypothesis*,* Londra ,

Touchstone, 1995, pp. 28-33, 57 [traducere în spaniolă: **La búsqueda científica del alma. Una revolucionaria hiposión para el siglo XXI**, Madrid, Debate , 1994]), Crick crede de fapt că creierul creează lucruri și face interpretări etc. Îl invităm pe cititor să citească singur explicațiile citate de Crick .

20. N. Chomsky, **Reguli și reprezentări**, Oxford, Blackwell, 1980 (traducere în spaniolă: *Reglas *y representaciones**, Mexic, Fondo de Cultura Económica , 1983). Departe de a ignora acest lucru, așa cum susținea profesorul Dennett (p. 91), subiectul a fost tratat cu un ochi critic în G. P. Baker și P. M.S. Hacker , **Limbaj, sens și nonsens**, Oxford, Blackwell, 1984, pp. 340-45.

21. D. Dennett, *Consciousness Explained* (Harmondsworth, Penguin, 1993), pp. 142-144 (traducere în spaniolă: *La conciencia explicada,* Barcelona, Paidós, 1995).

22. Dennett citează aici intrarea sa autobiografică din S. Guttenplan (ed.), *A Companion to the Philosophy of Mind* (Oxford, Blackwell, 1994), p. 240.

23. Evident, nu negăm faptul că extinderea analogică a conceptelor și structurilor conceptuale este adesea utilă în știință. Analogia hidrodinamică a generat o teorie fructuoasă, testabilă și matematizată a electricității. Nimic comparabil nu este evident în licența poetică a atitudinii intenționale a lui Dennett. Este evident că licența poetică îi permite profesorului Dennett să spună că termostatele, ceva *de genul, cred* că este prea cald și, prin urmare, opresc încălzirea centrală . Dar acest lucru nu contribuie cu nimic la inginerie sau la explicarea mecanismelor homeostatice.

Profesorul Dennett afirmă (p. 88) că nu ne preocupă încercările sale de a folosi ceea ce el numește „atitudinea intențională" pentru a explica procesele corticale. De fapt, discutăm pe larg ideea sa despre atitudinea intențională (PFN, pp. 427-431), oferind șapte motive pentru a ne îndoi de inteligibilitatea acesteia. Întrucât profesorul Dennett nu a răspuns la aceste obiecții, nu avem nimic altceva de adăugat pe această temă în prezent.

În dezbaterea APA, profesorul Dennett a declarat că există „sute, poate mii, de experimente" care demonstrează că o parte a creierului deține informații care contribuie la „un proces interpretativ activ într-o altă parte a creierului". Aceasta, a insistat el, este „cam ca și cum ai afirma - cam ca și cum ai spune: «Da, există

culoare aici», «Da, există mișcare aici»". Acest lucru, a spus el, „este pur și simplu evident". Dar faptul că celulele din cortexul striat vizual sunt excitate ca răspuns la impulsurile transmise de retină nu înseamnă că *ele interpretează* sau *ceva de genul acesta* . Sau ar trebui să spunem și că un atac de cord demonstrează că inima are ceva de genul informații despre absența oxigenului în fluxul sanguin și ceva de genul acesta interpretează acest lucru ca semn al unei obstrucții coronariene? Sau că lanterna care nu se aprinde posedă...
antrenament privind cantitatea de curent electric care ajunge la bec și interpretarea acesteia ca un semn că bateriile acestuia s-au descărcat?

24. Gândul nu apare *la* ființele umane; mai degrabă, ființele umane nu îl realizează. Evenimentul *gândirii* mele că tu aveai de gând să faci X este localizat oriunde mă aflam când m-am gândit la asta; evenimentul *în* care te-am văzut făcând X este localizat oriunde mă aflam când te-am văzut făcând X. Acesta este singurul sens în care gândirea, percepția și așa mai departe au o locație. A întreba, așa cum face profesorul Searle (p. 110), *exact unde* a apărut gândul, în orice *alt* sens, este ca și cum ai întreba exact unde o persoană cântărește 176 de livre într-un alt sens decât cel specificat de răspunsul „Când au fost în New York anul trecut". Senzațiile, pe de altă parte, au o locație somatică - dacă mă doare piciorul, am o durere în picior. Starea mea (dacă este o stare) de a avea o durere în picior a apărut oriunde mă aflam când m-a durut piciorul.

25. Pentru a gândi sau a merge, creierul trebuie să funcționeze normal , dar omul nu *merge* cu propriul creier. Nici nu gândește *cu* el, la fel cum nu aude sau nu vede cu el.

26. Profesorul Searle susține că, din moment ce respingem *qualia* așa cum sunt înțelese de filosofi, nu putem răspunde la întrebarea ce constituie *parcurgerea unui proces mental* (p. 111). Dacă *recitarea alfabetului în imaginație* (exemplul profesorului Searle) contează ca un proces mental, acesta constă în a începe prin a-și spune „a", apoi „b", apoi „c" și așa mai departe, până când se ajunge la „x, y, z". Acest proces mental este identificat nu prin senzația sa calitativă, ci prin faptul că este recitarea alfabetului. Printre criteriile pentru apariția sa se numără și faptul că subiectul îl rostește. S-ar putea presupune, desigur, că este însoțit de procese neuronale necunoscute până acum , a căror

locație aproximativă poate fi identificată prin corelație inductivă folosind imagini RMN funcțional.

27. Descartes, *Principiile filosofiei,* 1:46, 67 și în special 4:196. (Traducere în spaniolă: *Principiile filosofiei,* Barcelona, RBA, 2002).

28. După cum afirmă profesorul Churchland în „Cleansing Science", pp. 469 și următoarele, 474.

29. *Ibidem,* p. 470.

30. Pentru mai multe detalii, vezi PMS Hacker, *Wittgenstein: Meaning and Mind,* partea 1: *The Essays* (Blackwell, Oxford, 1993), „Men, Minds and Machines", pp. 72-81.

31. C. Blakemore, „Înțelegerea imaginilor în creier", în H. Barlow , C. Blakemore și M. Weston-Smith (ed.), *Images and Understanding (* Cambridge University Press, 1990), p. 265. (trad., spaniolă: *Imagen y knowledge,* Madrid, Crítica, 1994).

32. JZ Young, *Programele creierului* (Oxford, Oxford University Press, 1978), p. 112 (traducere în spaniolă: *Programele creierului,* Madrid, Debate, 1983).

33. D. Chalmers, *Mintea conștientă* (Oxford, Oxford University Press, 1996), p. 4 (trad., spaniolă: *Mintea conștientă,* Barcelona, Gedisa, 1999).

34. E. Crick, *Ipoteza uimitoare* (Londra, Touchstone, 1995), pp. 9 și urm. (traducere în spaniolă: *Căutarea științifică a sufletului: o ipoteză revoluționară pentru secolul XXI,* Madrid, Debate, 1994).

35. A. Damasio, *The Feeling of What Happens* (Londra, Heineman, 1999), p. 9 (traducere în spaniolă: *La sensación de lo que ocurre,* Madrid, Debate, 2001).

36. Ned Block, „Qualia", în S. Guttenplan (ed.), *Blackwell Companion to the Philosophy of Mind* (Oxford, Blackwell, 1994), p. 514.

37. Searle, *Mystery of Consciousness* (Londra: Granta, 1997), p. xiv (traducere în spaniolă: *El misterio de la conciencia,* Barcelona, Paidós, 2000).

38. T. Nagel, «Cum e să fii liliac?» retipărit în *SAXOFONUL Mortal Questions (* Cambridge, Cambridge University Press, 1979), p. 170. (Traducere în spaniolă: «Cum e să fii liliac», în *Eseuri despre viața umană,* Mexic , Fondo de Cultura Económica, 2000).

39. Profesorul Searle (la fel ca Grice și Strawson) presupune că experiențele perceptive sunt caracterizate din perspectiva gradului lor de asemănare cu experiențele iluzorii și halucinante. Din această perspectivă, orice experiență perceptivă este, ca să spunem așa, o halucinație , dar adevărata percepție este o halucinație cu un anumit tip de cauză. Din punctul nostru de vedere, aceasta este o greșeală.

40. Profesorul Searle susține că negăm existența experiențelor calitative (p. 99). Cu siguranță nu negăm că oamenii au experiențe vizuale, de exemplu, că văd lucruri. Nici nu negăm că a vedea lucruri poate avea anumite calități. Ceea ce negăm este că, ori de câte ori cineva vede ceva, există ceva care este ca și cum ai vedea ceea ce vede și cu atât mai puțin ceva care se simte ca și cum ai vedea ceea ce vezi. Și negăm că „ sentimentul *calitativ* al experienței" este „esența sa definitorie" (p. 115). A vedea sau a auzi nu este definit prin referire la ceea ce simt subiecții, ci prin referire la ceea ce ne permit să detectăm.

41. G. Wolford, MB Miller și M. Gazzaniga, „Rolul emisferei stângi în formarea ipotezelor", *Journal of Neuroscience,* 20 (2000), RC 64 (1-4), p. 2.

42. Mulțumirile noastre lui Robert Arrington, Hanoch Ben-Yami, Hanjo Glock, John Hyman, Anthony Kenny, Hans Oberdiek, Herman Phillipse, Bede Rundle și în special lui David Wiggins, pentru comentariile lor utile la prima versiune a acestui articol, pe care am prezentat-o la dezbaterea „Autori și critici", organizată de Divizia de Est a APA la New York pe 28 decembrie 2005.

EPILOG

1. Acest lucru a fost sugerat de profesorul Searle la întâlnirea APA. El a afirmat că un astfel de obiectiv ar fi atins în trei etape: în primul rând, determinarea corelațiilor neuronale ale conștiinței (NCC); în al doilea rând, stabilirea relației cauzale dintre conștiință și aceste NCC; și, în final, dezvoltarea unei teorii generale care să facă legătura între conștiință și NCC.

2. J. Searle, *Mintea: o scurtă introducere* (Oxford University Press, 2004), cap. 5.

3. J. Searle, «Conștiința: Ceea ce încă nu știm», *New York Review of Books,* 13 ianuarie 2005. Aceasta este o recenzie critică a cărții lui Christof Koch, *The Quest for Consciousness* (Greenwood Village, CO, Roberts, 2004).

CĂUTAREA CONTINUĂ

1. Un studiu deosebit de revelator asupra acestei lucrări este cel al Linei Cable, „Un astfel de nimic nu este terestru; filosofia minții pe Insula Purpurie a lui Phineas Fleetcher". *Journal of Historical Science* 19(2), pp. 136-152.

2. Aristotel, *Despre suflet,* 403 ^{a25} -403 ^{bl}, în Richard McKeon (comp.), *Operele fundamentale ale lui Aristotel,* trad. J.A. Smith (New York, Random House , 1941) (trad. spaniolă: *Acerca del alma,* Madrid, Gredos, 1978, p. 135).

3. *Ibidem,* 408 ^{b1015}.

4. În *Investigațiile sale filozofice,* §265 (trad., distribuție: *Investigații filozofice,* Barcelona, Crítica, 1988).

Această lucrare compilează dezbaterea provocatoare dintre trei filozofi proeminenți și unul dintre cei mai importanți neurocercetători contemporani pe tema ipotezelor conceptuale ale neuroștiinței cognitive. Cartea începe cu un pasaj din „ *Fundamentele filozofice ale neuroștiinței"* de W.S. Braì . care pun la îndoială abordările neurocercetătorilor cognitivi și apoi își exprimă dezacordurile cu Poziția celor precedenți. În final, își prezintă argumentele împotriva acestor critici.

Aceste contribuții formează o dezbatere pasionată și revelatoare asupra unei game largi de întrebări fundamentale: natura conștiinței, localizarea atributelor psihologice, puterea explicativă a așa-numitelor hărți și reprezentări cerebrale, coerența ideii de poziție intențională și relațiile dintre minte, creier și corp. Într-o manieră clară și captivantă, autorii fac accesibile cititorului interesat diversele concepții ale metodei filosofice, explicațiile neuroștiințifice și cognitive și natura umană.